Voluntary Disclosure of Information and Cooperation in Simultaneous-Move Economic Interactions

Kamei, Kenju

Durham University

3 January 2020

Voluntary Disclosure of Information and Cooperation

in Simultaneous-Move Economic Interactions

Kenju Kamei*

Department of Economics and Finance, University of Durham,

Email: kenju.kamei@gmail.com, kenju.kamei@durham.ac.uk.

Abstract: This paper experimentally studies individuals' voluntary disclosure of past behaviors and its effects in a finitely repeated two-player public goods game. The experiment data found that voluntary information disclosure strengthens cooperation under certain conditions, although a non-negligible fraction of individuals do not disclose information about the past and proceed to behave opportunistically. On closer inspection, the data revealed that the material incentives of disclosure acts differ according to the matching protocol. Specifically, disclosers receive higher payoffs than non-disclosers if the disclosers are assured to be matched with like-minded disclosers; conversely, disclosers are vulnerable to exploitation by others under random matching. These results suggest that mandatory disclosure helps enhance economic efficiency if individuals' hiding and uncooperative behaviors are liable to precipitate a collapse in the community norms.

Keywords: experiment, information disclosure, cooperation, dilemma, repeated games, reputation

JEL code: C92, D74, D83

* This project was supported by research grants from the Telecommunications Advancement Foundation and from the Murata Science Foundation. Durham University Business School provided additional funding. The author thanks the IT team at the Durham University Business School and Paudie Lynch (the IT manager) for help in managing the ORSEE recruiting system and computers when he conducted the experiment, and Burkhard Schipper, Matthew Walker, Louis Putterman, and audience in 2018 BABEEW (in San Francisco) and in Hitotsubashi University for helpful comments. The author also thanks the editor, Daniela Puzzello, and two anonymous referees for valuable comments. This paper was previously circulated with a different title: "Information Disclosure and Cooperation in a Finitely-repeated Dilemma: Experimental Evidence."

1. Introduction

This paper experimentally explores whether voluntary disclosure of past behaviors improves cooperation in a public goods game. Answering this question helps understand recent development of reputation mechanisms on online-based transactions (e.g., Uber). The reputation mechanism in the field is a forced disclosure system, whereby users' ratings and reviews for past transactions are *always* available to potential counterparties. Prior experiments have shown that the mandatory disclosure of information on individuals' past behaviors can mitigate dilemmas and improve cooperation, even in finitely repeated setups (e.g., Engelmann and Fischbacher, 2009; Bolton *et al.*, 2004, 2005; Kamei and Putterman, 2017).[1] While economists have recently begun to study people's possible endogenous formation of reputational information in several formats,[2] it remains unclear how the market would look if users could *voluntarily* disclose their reputations (past behaviors). A voluntary disclosure system may be theoretically similar to a mandatory disclosure system if 'rational' cooperation occurs as discussed in Kreps *et al.* (1982). But if this is the case, one might wonder why mandatory, not voluntary, disclosure has been consistently adopted by the online platforms for users' ratings.[3]

This study utilizes the framework of a finitely repeated two-player public goods game for its sharp standard theory prediction. A parsimonious design is adopted by using a simple disclosure format. If a player decides to disclose in a given period, her matched partner observes the player's last period action choice before their interaction commences. A similar, but exogenous, disclosure system has been used in prior prisoner's dilemma game setups (e.g., Kandori, 1992; Stahl, 2013) and investment game setups (e.g., Bolton *et al.*, 2004).

---

[1] The positive impact of exogenously given information has also been demonstrated in indefinitely repeated dilemma game experiments with random matching (e.g., Camera and Casari, 2009; Kamei, 2017).
[2] Recent research explored people's reputation building through gossiping (e.g., Camera and Casari, 2018; Kamei and Putterman, 2018), information acquisition (e.g., Camera and Casari, 2018; Duffy *et al.*, 2013), identity disclosure (Kamei, 2017), and endogenous dissolution of relationships (e.g., Honhon and Hyndman, 2019).
[3] Although mandatory disclosure is more common than voluntary disclosure, it is not universal. For instance, in China, people hold reputational information built on online platforms, so-called 'social credit', and they are actively encouraged to reveal the scores to potential transaction counterparts even outside the platforms (e.g., Hatton, 2015). Users' disclosure behaviors and the value of disclosure are of great interest to scholars as well as practitioners.

The results of the experiment show that endogenously disclosed information improves cooperation under certain conditions. However, not everyone discloses, even if disclosure is free. Besides, the impact of disclosed information is weak when subjects are randomly matched with other subjects in their community in each period (especially when disclosure involves a cost) because then a large fraction of subjects are reluctant to disclose their past. A control treatment with mandatory disclosure reveals that subjects are able to sustain high cooperation norms if all subjects are *forced* to reveal their past actions. In sum, mandatory disclosure can be a more efficient mechanism than voluntary disclosure for reviews on online platforms.

The experiment further shows that sorting based on disclosure decisions helps disclosers achieve a high degree of cooperation with other disclosers.[4] The level of cooperation observed among the disclosers is comparable to that under mandatory information disclosure. Nevertheless, even with sorting, a non-negligible fraction of subjects remain non-disclosers, contribute small amounts, and keep receiving low payoffs as a result.

The closest paper to the present one is Kamei (2017), who focused on people's reputation building behaviors through identity disclosure in an indefinitely repeated prisoner's dilemma game where mutual cooperation holds as an equilibrium outcome. He found that almost everyone discloses their identity and successfully cooperates with each other when hiding it is costly (40% of the deviation gain). Aside from the time horizon (finitely versus infinitely repeated), the present paper differs from Kamei (2017) in three important aspects. First, this study retains anonymity to isolate the effect of disclosed information, showing that the disclosed information of own action choices per se has a positive effect.[5] Second, this study examines the effect of a disclosure cost

[4] There are many real-world situations where individuals' sorting behavior is linked to their information disclosure. For example, in online markets, individuals can usually access platforms that do not have reputation mechanisms in addition to ones with mandatory disclosure (e.g., dating services), and they need to decide which platform to join.
[5] It is worth noting that this paper is also related to the literature on the voluntary information disclosure under asymmetric information (e.g., Dranove and Jin [2010]). Using theoretical models and empirical data, past studies examined cases in which one party (seller) possesses private information on an exogenous state (e.g., product quality), and the other party (buyer) chooses an action after observing both the seller's disclosure decision and the contents of disclosed information. The voluntary information disclosure in this context also has received scant attention in the literature on laboratory experiments to date. Benndorf *et al.* (2015) remark that: "only Forsythe *et al.* (1989) have

rather than a hiding cost, and sets the cost at only 5% of the per-period endowment. The new experiment finds that such a small cost strongly affects the subjects' behaviors. Third, this study compares reputation building behaviors between two matching protocols: random matching versus sorting, showing that individuals' contribution behaviors (including material benefits of disclosing acts) and the impact of the disclosure cost differ drastically according to the matching protocol. Kamei (2017), in contrast, studied people's identity disclosure behaviors only under random matching.

The rest of the paper proceeds as follows: Section 2 summarizes the experimental design, Section 3 theoretically discusses subjects' possible behaviors, Sections 4 and 5 report the experimental results, and Section 6 concludes.

2. Experimental Design

This study is built on the framework of a two-player linear public goods game. The experiment consists of 20 interaction periods. The number of interactions is common knowledge to all subjects. In each period, every subject is paired with another subject, is given an endowment of 20 ECUs (experimental currency units) and simultaneously decides how many ECUs they wish to contribute to their pair's joint account. Subjects' contribution amounts must be integers between 0 and 20. The marginal per capita return (MPCR) is 0.8. Thus, the payoff of subject i ($\pi_{i,t}$) is calculated as follows:

$$\pi_{i,t} = 20 - c_{i,t} + .8 \cdot (c_{i,t} + c_{j,t}), \qquad (1)$$

where $c_{i,t}$ is subject i's contribution to her joint account in period t, and subject j is i's matched person in period t. In order to study the pure impact of voluntary information disclosure in a controlled manner, no subject identification numbers are provided in any session.[6]

studied unraveling in an experiment." Three further experiments have recently studied subjects' voluntary disclosure in sequential transactions with exogenous states (Hagenbach and Perez-Richet, 2018; Jin *et al.*, 2015; Li and Schipper, 2018). Unlike all prior studies, Kamei (2017) and this paper study voluntary information disclosure, where the states are *endogenous* (determined by prior actions).

[6] As an anonymous referee pointed out, if a subject i is matched with j in two consecutive periods (periods t and t + 1) and j discloses his allocation amount in period t + 1, i may assume that i was matched with the same person j in both periods t and t + 1 because of the rich choice space, {0, 1, 2, …, 20}, in the contribution decisions (see

This study consists of four main treatments (Section 2.1) and control treatments (Section 2.2).

2.1. Four Main Treatments

In period 1, each subject is randomly matched with another subject without making any disclosure decisions; and then plays the two-player public goods game in their matched pair. Each period *t*, where $t > 1$, consists of two stages. In the first stage, subjects decide whether to disclose their period $t - 1$ contribution decisions to their period *t* partners. When a subject chooses to disclose, her period *t* partner is informed of the previous contribution decision; the partner is not given this information when the subject decides not to disclose.

This study constructs four treatments by varying two dimensions using a 2×2 between-subjects design (Table 1). The first dimension is the size of the disclosure cost: either the disclosure is free or costs one ECU.[7] In costly disclosing treatments, one ECU (equal to 5% of the endowment) is deducted from a discloser at the end of a given period (a subject has 20 ECUs in her allocation-decision stage even when she decides to disclose). The second dimension is the matching protocol: either random matching (each subject is randomly matched with another, regardless of the disclosure decision) or sorting (each discloser [non-discloser] is randomly matched with another discloser [non-discloser] in a given experimental session).[8] The four treatments are denoted as "Costly Disclosure, Random Matching" (C-RM), "Free Disclosure, Random Matching" (F-RM), "Costly Sorting" (C-Sorting), and "Free Sorting" (F-Sorting) treatments.

Appendix Fig. C.2 for the cumulative distributions of individual contribution amounts). This feature may provide subjects some incentive to build a cooperative reputation. However, any impact should be small for two reasons. First, since the session size is 12 in most sessions (footnote 14), the re-matching probability is low (e.g., less than 10% under random matching). Second, Kamei and Putterman (2017) ran the treatments with a finitely repeated two-player linear public goods game where only subjects' IDs were available to each other (the LI-HG and LI-LG treatments). Despite being given an ability to choose their partners, cooperation did not evolve in these treatments.

[7] Disclosure is costly in some real-world situations. For example, on some online dating service, singles may disclose their detailed background information by spending time filling in registration forms and/or updating their profiles, and also paying membership fees.

[8] If the number of disclosers is an odd number in a given session under the sorting condition, one discloser is randomly matched with a non-discloser. This event happened in only 9.3% (6.9%) of pairings in the C-Sorting (F-Sorting) treatment. The paper's findings are robust regardless of whether only data consisting of pairs with the same preferences (two disclosers or two non-disclosers) or all data are used.

Table 1. *Summary of Treatments*

Treatment name	Disclosure cost	Matching protocol	# of sessions (subjects)
[Main treatments:]			
C-RM	1 ECU	Random Matching	4 (48)
C-Sorting	1 ECU	Sorting	4 (44)
F-RM	0 ECUs	Random Matching	4 (44)
F-Sorting	0 ECUs	Sorting	4 (44)
[Control treatments:]			
Baseline	n.a.	Random Matching	4 (44)
C-Sorting-N	1 ECU	Sorting	4 (44)
Mandatory	n.a.	Random Matching	2 (24)
Total			26 (292)

Notes: Random Matching = Each subject is randomly matched with another subject in a session. Sorting = Each discloser (non-discloser) is matched with another discloser (non-discloser) in a session.

2.1.1. Conditional Contribution Schedule and Belief Elicitation

Two additional tasks are included. The first task is the elicitation of beliefs. In each allocation-decision stage, subjects are asked about their beliefs regarding the matched partner's contribution amount in a given period *t*.[9] Subjects in the C-RM and F-RM treatments are also asked to state their expectation as to the number of disclosers (except themselves) in period *t*. As the primary focus is on subjects' behaviors and because incentivized elicitation may affect subjects' actual behaviors (e.g., Gächter and Renner 2010), these tasks were not incentivized.[10,11]

Second, cooperation types are elicited from subjects using the method of Fischbacher *et al.* (2001). Specifically, each subject is asked how many ECUs, given an endowment of 20 ECUs, they wish to allocate to their group, conditional on each of the other group members' average

[9] In the elicitation stage, subjects are aware of their partners' current-period disclosure decisions and also the partners' last-period contribution amounts in the case that the partners selected to disclose.
[10] These elicitation tasks were only explained in the instructions shown on subjects' computer screens, not in the hard copy of instructions distributed to subjects, in order to avoid making the tasks salient.
[11] Gächter and Renner (2010) showed that incentivized elicitation improves belief accuracy only a little, while it significantly influences subjects' decisions to contribute in a public goods game. One possible way to incentivize beliefs is to randomly select some periods for payments based on belief accuracy and the other for payments based on the actual contribution behaviors. To reduce the design complexity, this method was not employed.

contributions (= 0, 1, 2, …, 20). Classified types based on the conditional schedules are used to study how subjects' disclosure and contribution decisions differ by their intrinsic propensity to cooperate. This task is incentivized and is included before the finitely repeated public goods game.[12] Subjects are, however, informed of the outcomes of this task only after they complete the 20 periods of the public goods game. In addition, neither the group composition nor the outcome affects the main repeated dilemma interactions (e.g., pairing process).

2.2. Control Treatments

Three control treatments were additionally conducted (Table 1). First, in the "Baseline" treatment, subjects are *not* allowed to disclose their states, are just randomly matched with another subject in a session, and play the public goods game in each period. This treatment serves as a control to identify the impact of information disclosure and/or sorting in the main treatments.

Second, the "Mandatory" ("Mandatory Disclosure") treatment will be used to identify the impact of *exogenous* information disclosure and to compare it against the impact of voluntary information disclosure in the C-Sorting treatment – the treatment which displayed the highest efficiency among the four main treatments (Section 4). At the onset of the Mandatory treatment, subjects are randomly assigned to a group of six and the group composition stays the same during the experiment.[13] In each period, subjects are randomly matched with one another in their own group and play the two-player public goods game (Equation (1)). From period 2, each subject's last period contribution amount is always automatically revealed to their partner.

Lastly, the third control treatment is called the "C-Sorting-N" (Costly Sorting, No Disclosure) treatment. This treatment is identical to the C-Sorting treatment, except that the

---

[12] In this task, subjects are randomly assigned into groups of four and make two kinds of decisions in a linear public goods game with an MPCR of 0.4. The first decision is the conditional contribution decisions, as just explained. The second decision is unconditional contribution decisions (i.e., each subject decides how much to contribute to their group unconditionally). Once all subjects have made the two decisions, one subject is randomly selected as the one whose conditional schedule is used to calculate her contribution amount. For the remaining three subjects in the group, their unconditional contribution decisions are used for their contribution amounts (see online Appendix A).

[13] This setup was selected because, as will be explained in Sections 4 and 5, two subgroups – one for disclosers and the other for non-disclosers – were formed, and the average subgroup size was six in the C-Sorting treatment (Fig. 1). In addition, subjects' mobility between the two subgroups was small in the C-Sorting treatment (Section 5.3).

information is not disclosed. Subjects experience two stages after period 1. In the first stage, they decide whether or not to pay one ECU. A player who pays (does not pay) the fee is called the "payer" ("non-payer"). Payers (Non-payers) are randomly matched with another payer (non-payer). However, their last-period contributions are *not* revealed. The data for this treatment are used to study the relative importance of (a) the presence of a matching cost and (b) disclosed information in determining high performance in the C-Sorting treatment. Since this treatment is not directly related to the hypotheses of the paper, the results are relegated to Appendix E.

2.3. Experimental Procedure

A total of 26 sessions (16 for the four main treatments and ten for the control treatments) were conducted from August 2015 through August 2017. All participants were Durham University students. No students participated in more than one session. All the instructions were neutrally framed. Communication was prohibited during the entire experiment. Subjects were privately paid based on their accumulated ECUs (40 ECUs were exchanged for £1) at the end of the experiment.[14] The average payment (including a show-up fee of £3) was £15.61. The average experiment duration (including reading instructions and paying subjects) was 90 minutes.

3. Discussions on Subjects' Possible Disclosure and Reputation Building

Standard theory, based on players' self-interest and common knowledge of rationality, provides a point prediction because the MPCR is 0.8. Contributing zero to the joint account is a dominant strategy for each player in any period ($\partial\pi_{i,t}/\partial c_{i,t} < 0$). Thus, by the logic of backward induction, each player would contribute nothing in every period under the assumption that they believe their peers would always contribute zero. Considering the peers' full and uniform free-riding behavior, no one would incur a cost to disclose their past in any period in the C-RM and C-Sorting treatments. Disclosure decisions do not affect players' payoffs in the F-RM and F-Sorting

[14] The session size was 12 in all sessions, except for one session in each of the C-Sorting, F-RM, F-Sorting, Baseline and C-Sorting-N treatments. For these sessions, the session size was eight.

treatments as these actions can be taken for free and the peers would always contribute zero; hence, players would randomly decide whether to disclose in these two treatments.

Hypothesis 1: *Standard theory prediction*.

(a) No one discloses their state (last-period contribution amount) in the C-RM and C-Sorting treatments. (b) Disclosure decisions are randomly made in the F-RM and F-Sorting treatments. (c) Subjects contribute nothing to the joint accounts in each period in all treatments.

In the absence of institutions to assist people's contribution behaviors, a large body of experimental research partially supports Hypothesis 1(c). For example, although in earlier periods subjects may contribute around 40% to 60% of their endowment in a finitely repeated public goods game, they decrease their levels of cooperation steadily over time (Ledyard, 1995; Chaudhuri, 2011).[15] With voluntary information disclosure, however, cooperation could evolve, similar to the logic of Kreps *et al.* (1982), if we assume that there exist players who choose to disclose and then act according to a conditional cooperation strategy (i.e., contribute conditional on partners doing so). The prevalence of conditional cooperators is well-documented (e.g., Fischbacher *et al.*, 2001; Fischbacher and Gächter, 2010). Provided that there is a sufficiently large fraction of subjects who act following the conditional cooperation strategy, selfish free rider type i has incentives to *strategically* mimic the behavior of a conditional cooperator, and Hypothesis 2 can be derived (see Appendix B for the mathematical detail). Here, $D(.)$ and $C(.)$ refer to the disclosure rate and average contribution amount, respectively, in a given treatment.

Hypothesis 2: *Rational cooperation in the presence of conditional cooperators*

(a) Sorting improves cooperation: D(C-Sorting) > D(C-RM) and D(F-Sorting) > D(F-RM). Likewise, C(C-Sorting) > C(C-RM) and C(F-Sorting) > C(F-RM).

(b) A disclosure cost does not discourage cooperation under random matching: C(F-RM) = C(C-RM) while D(F-RM) ≥ D(C-RM).

[15] Similar free-riding dynamics are reported when the group size is two, irrespective of whether partner or random matching is used (e.g., Kamei, 2019).

(c) Mandatory disclosure improves cooperation: C(Mandatory) > C(C-RM) = C(F-RM).

An intuition behind Hypothesis 2(a) is that sorting increases the likelihood that free rider types interact with conditional cooperators if the former mimic the behavior of the latter. Past experiments suggest that endogenous regrouping encourages selfish types to strategically mimic cooperative types and achieve high efficiency (e.g., Page *et al.*, 2005; Kamei and Putterman, 2017).

Hypothesis 2(b) states that theoretically, the disclosure cost does not affect subjects' contribution behavior under random matching, as the cost is too small: just one ECU. It is worth noting that this may not be correct behaviorally if subjects have a discontinuity in their cooperation tendency between free versus costly disclosure (e.g., Shampanier *et al.*, 2007).

Lastly, Kamei (2017) supports Hypothesis 2(c). Kamei (2017), in the context of a prisoner's dilemma game, showed that: (i) given an option to hide IDs, a non-negligible fraction of individuals do not disclose their IDs and behave uncooperatively, even if mutual cooperation holds as an equilibrium outcome; and (ii) some subjects do not learn to disclose, even though disclosers continue to select defection against masked subjects. Such harmful hiding and uncooperative behaviors might be even stronger in the present study due to the finitely repeated setup.[16]

4. Treatment Effects of Disclosure and Better Matching

An overview of subjects' decisions is given in Section 4.1. The treatment differences are examined in Section 4.2 and the material benefits of disclosing acts are investigated in Section 4.3.

4.1. Overview of the Experiment

First, the effect of exogenously disclosed information can be studied using the data of the Baseline and Mandatory treatments. Fig. 1 shows subjects' average behaviors while dividing the

[16] No clear prediction is possible for a comparison between the sorting versus the Mandatory treatments. The percentage of mimickers may be larger in the sorting than in the Mandatory treatment, considering that sorting assures a better matching among like-minded individuals. However, the opposite may also be possible if eliminating the option to hide encourages those who would hide and then behave uncooperatively in the sorting treatments to strategically build a reputation under mandatory disclosure.

data into the first half (periods 2-10) and the second half (periods 11-20) of the experiment.[17,18] It suggests that groups achieve high efficiency if everyone is forced to disclose their state (panels (a) and (c)). The overall average contribution in the Mandatory treatment was about three times as large as in the Baseline treatment. A similar difference was also observed for subjects' beliefs (panel (b)). The strong effect of mandatory disclosure on improving cooperation is consistent with the prior studies that found positive information effects in the contexts of a helping game (e.g., Bolton *et al.*, 2005) and an investment game (e.g., Bolton *et al.*, 2004).

Fig. 1 further includes the data of the four main treatments. The majority of Hypothesis 2 hold, although it does not perfectly explain the subjects' behaviors. First, a non-negligible fraction of subjects chose to disclose their states, irrespective of whether disclosure was costly or free (see panel (d)).[19] Consistent with Hypothesis 2(a), the disclosure rate (the percentage of the cases in which subjects chose to disclose) was higher with sorting than under random matching for a given disclosure cost, although it remained well below 100% even with sorting. Further, the presence of a positive disclosure cost undermined subjects' disclosure rates.

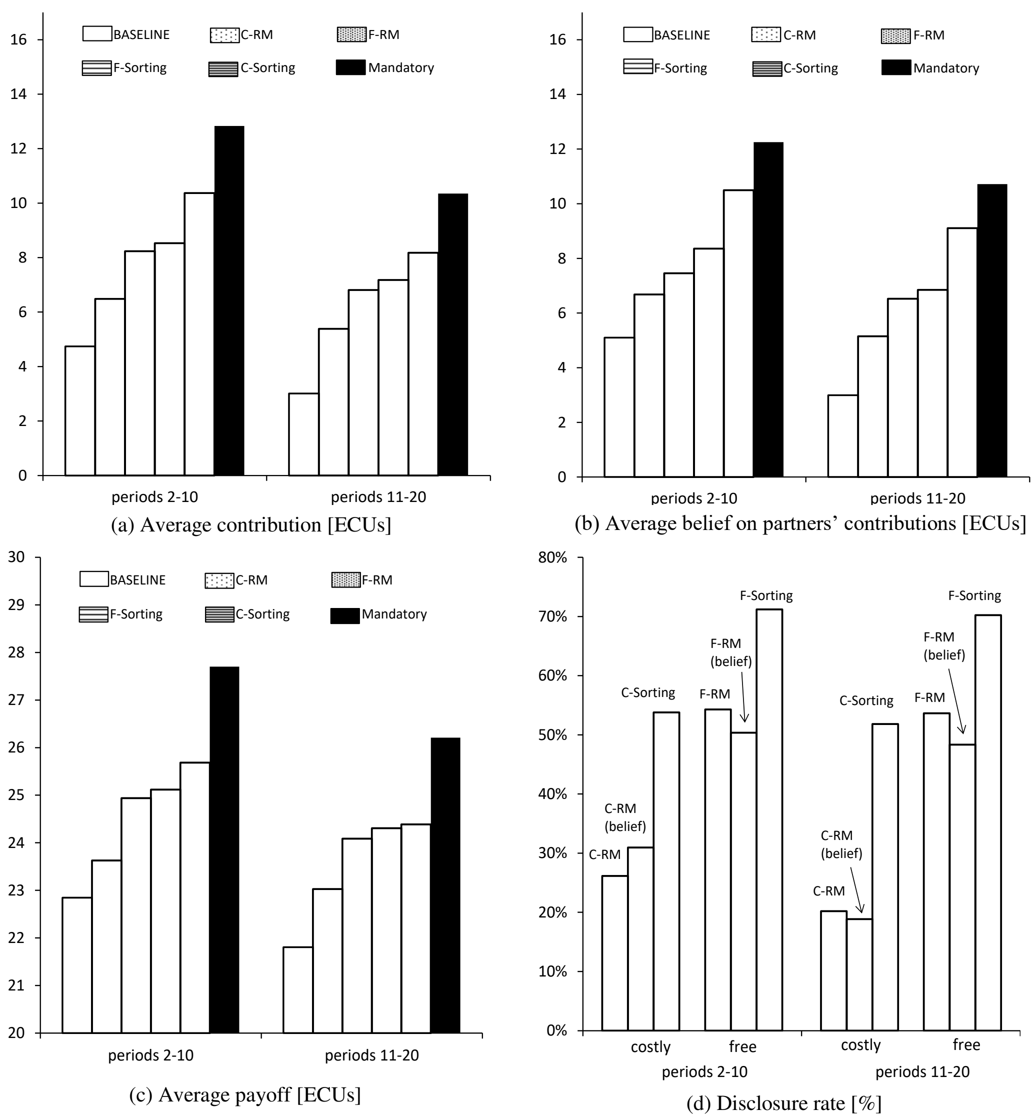
Second, voluntary information disclosure has a positive effect on improving cooperation, and accordingly payoff, in all the four treatments (panels (a) and (c)). As predicted by Hypothesis 2(a), the positive effect was stronger with sorting than under random matching for a given disclosure cost. However, the difference between the F-RM and F-Sorting treatments was small.

Third, the positive disclosure cost helped improve cooperation with sorting. This implies that a higher disclosure rate due to the lack of a disclosure cost is not helpful in improving cooperation. By clear contrast, the positive disclosure cost strongly undermined cooperation under

[17] The data from period 1 were excluded because subjects were not given an opportunity to disclose the state in that period in the treatments with information disclosure. Results are nevertheless similar even if all data including period 1 are used (the results are omitted to conserve space).
[18] Fig. 1 indicates that subjects' average contributions, beliefs and payoffs were all on average lower in the second half than in the first half of every treatment, including these two control treatments. This is a natural end-game effect (Andreoni, 1988). The treatment differences were similarly observed regardless of which half of the data are used for a comparison. See Appendix Fig. C.1 for subjects' period-by-period decisions.
[19] Subjects correctly anticipated the peers' disclosure behaviors in the C-RM and F-RM treatments (panel (d)).

random matching, unlike Hypothesis 2(b). The cost may have driven some cooperative types to non-disclosure and conservative behavior. This interpretation is consistent with the low disclosure rate result in the C-RM treatment (only 23.0% on average), which decreases the material incentives to mimic the behavior of a conditional cooperator. This implies that subjects may have a discontinuity in their reputation building behaviors between positive and zero costs (e.g., Shampanier *et al.*, 2007).

Fig. 1: *Subjects' Average Behaviors in the First and Second Halves of the Experiment*



(a) Average contribution [ECUs]

(b) Average belief on partners' contributions [ECUs]

(c) Average payoff [ECUs]

(d) Disclosure rate [%]

Notes: The average payoffs of the Baseline, Mandatory, F-RM and F-Sorting treatments are the monotonic transformation of average contributions based on Equation (1). The average payoffs of the C-RM and C-Sorting treatments are the monotonic transformation of average contributions, minus the average disclosure costs.

Subjects' possible overreaction to the positive disclosure cost under random matching is also evident from the dynamics of their disclosure decisions. The subjects' disclosure rate displayed a decreasing trend in the C-RM treatment, while it remained stable in the F-RM, C-Sorting and F-Sorting treatments (Appendix Fig. C.1.I(c) and II(c), Fig. 1(d)). The trend in the C-RM treatment may have occurred because disclosure involved a cost and no sorting mechanism was present. Thus, subjects might have gradually perceived that disclosure was not worth the cost.

Lastly, the efficiency under mandatory disclosure was the highest in this study.

4.2. Disclosure Behaviors and the Impact of Disclosed Information

The following studies the treatment differences using session-average observations (Table 2).[20] Consistent with the discussions in Section 4.1, the disclosure rates were significantly higher in the C-Sorting (F-Sorting) than in the C-RM (F-RM) treatment.[21] In addition, the presence of a positive disclosure cost strongly influenced subjects' decisions to disclose: the disclosure rates were significantly lower in the C-RM (C-Sorting) than in the F-RM (F-Sorting) treatment.

Result 1: *(i) Non-negligible fractions of subjects disclosed their states in all conditions. (ii) Subjects' disclosure rates were significantly higher with sorting than under random matching for a given disclosure cost (whether costly or free). (iii) A positive disclosure cost significantly decreased subjects' disclosure rates, whether random matching or sorting was used.*

The calculations also show that the impact of information disclosure on efficiency depends on the matching protocols. The impact is significant in the C-Sorting and F-Sorting treatments, but not in the C-RM and F-RM treatments (columns (2) and (4)).[22] The effect of sorting, nevertheless, differs by the disclosure cost. First, the average contribution and payoff were both significantly larger in the C-Sorting than in the C-RM treatment. However, sorting had little impact on

[20] The treatment differences were also studied based on a regression model using session-average observations. The results are similar, but have more significant differences (see Appendix Table C.1 for the detail).
[21] Subjects' disclosure rates were significantly positive in all four treatments (column (1) of Appendix Table C.1).
[22] The average contribution in the F-RM treatment is about double that in the Baseline treatment. However, the difference in the average contribution is not significant due to a high variance in the former treatment. The difference is significant if a regression method is used (column (2) of Appendix Table C.1).

efficiency when disclosure was free, despite Result 1(ii). For example, the average contribution was 7.48 ECUs in the F-RM treatment vs. 7.82 ECUs in the F-Sorting treatment.

Consistent with Hypothesis 2(c), Table 2 also confirms the outperformance of mandatory disclosure under random matching. The efficiency, whether in contributions or payoffs, in the Mandatory treatment was significantly (weaker significantly) higher than in the C-RM (F-RM) treatment. Although the differences are not significant, the average contribution amounts were 25.2% and 47.4% higher in the Mandatory treatment than in the C-Sorting and F-Sorting treatments, respectively.[23]

Table 2. *Treatment Differences in Subjects' Behaviors*

	Disclosure rates (1)	Avg. contributions (2)	Avg. beliefs (3)	Avg. payoffs (4)
A. Average subjects' behaviors by treatment:				
(i) Baseline treatment	n.a.	3.91 [19.6%]	3.99 [20.0%]	22.30
(ii) C-RM treatment	23.0%	5.90 [29.5%]	5.87 [29.4%]	23.31
(iii) C-Sorting treatment	52.8%	9.21 [46.1%]	9.76 [48.8%]	25.00
(iv) F-RM treatment	53.9%	7.48 [37.4%]	6.97 [34.8%]	24.49
(v) F-Sorting treatment	70.7%	7.82 [39.1%]	7.56 [37.8%]	24.69
(vi) Mandatory treatment	n.a.	11.53 [57.7%]	11.45 [57.3%]	26.92
B. *p*-value (two-sided) for Mann-Whitney tests based session-average observations:				
H_0: (i) = (ii)	n.a.	.3865	.2482	.3865
H_0: (i) = (iii)	n.a.	.0209**	.0209**	.0833*
H_0: (i) = (iv)	n.a.	.1489	.1489	.1489
H_0: (i) = (v)	n.a.	.0433**	.0833*	.0433**
H_0: (i) = (iv)	n.a.	.0433**	.0209**	.0433**
H_0: (ii) = (iii)	.0209**	.0209**	.0209**	.0209**
H_0: (ii) = (iv)	.0209**	.2482	.5637	.1489
H_0: (ii) = (v)	.0209**	.1489	.1489	.1489
H_0: (ii) = (vi)	n.a.	.0209**	.0433**	.0209**
H_0: (iii) = (iv)	.5637	.1489	.0833*	.5637
H_0: (iii) = (v)	.0433**	.2482	.0433**	.7728
H_0: (iii) = (vi)	n.a.	.2482	.5637	.1489
H_0: (iv) = (v)	.0433**	.7728	.7728	.7728
H_0: (iv) = (vi)	n.a.	.0833*	.0833*	.0833*
H_0: (v) = (vi)	n.a.	.1489	.1489	.1489

Notes: Averaged from period 2 to 20. The numbers in squared bracket in column (2) (column (3)) are average contributions (beliefs) as the percentages of endowments. Results are similar when the averages are calculated using all data including the one from period 1. *, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

[23] According to a regression analysis, the average contribution and payoff are both weakly significantly higher in the Mandatory than in the F-Sorting treatment. The average payoff is also weakly significantly higher in the Mandatory compared to in the C-Sorting treatment (Appendix Table C.1).

Result 2: *(i) Information disclosure significantly improved efficiency (whether in contributions or payoffs) under sorting, but not under random matching. (ii) The average contribution and payoff were both significantly higher in the C-Sorting than in the C-RM treatment. (iii) Mandatory disclosure strongly improved efficiency.*

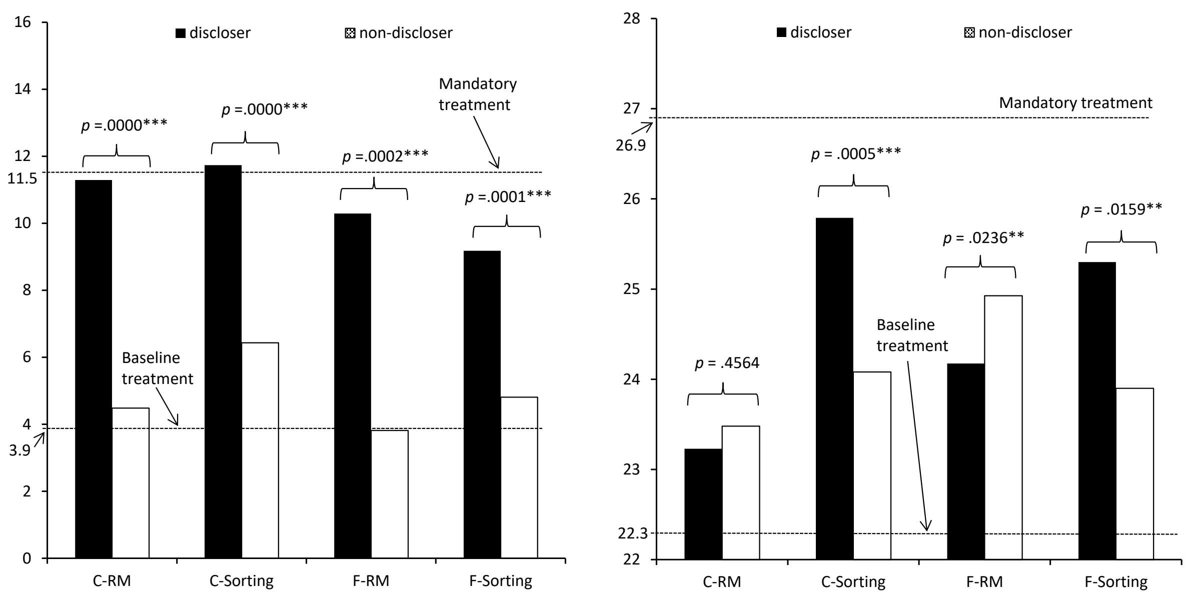
4.3. Subjects' Benefits of Disclosing Acts

Theoretically, free riders have strategic incentives to build cooperative reputations under certain conditions (Section 3 and Appendix B). A close look at the data by disclosure decision reveals that disclosers contributed more strongly than non-disclosers in all four treatments (Fig. 2(a)).[24] However, this does not mean that subjects had incentives to build cooperative reputations in all four treatments. As shown in Fig. 2(b), there is a clear contrast regarding the material consequences of the disclosing acts between the matching protocols. In the C-RM and F-RM treatments, disclosers received on average *lower* payoffs than non-disclosers. The difference is significant for the F-RM treatment.[25] This suggests that under random matching, disclosers' intentions to cooperate or build reputations were exploited by non-disclosers. By sharp contrast, in the C-Sorting and F-Sorting treatments, disclosers received on average *higher* payoffs than non-disclosers.[26,27] This underscores the role of sorting in making information disclosure materially beneficial and is consistent with the treatment differences summarized in Result 2(i).[28]

[24] This holds regardless of whether the data only from the first half (periods 2 to 10) or from the second half (periods 11 to 20) are used. See Appendix Fig. C.4.
[25] The lack of significance in the C-RM treatment is due to subjects' low disclosure rates in the C-RM treatment. Non-disclosers were less frequently matched with disclosers in the C-RM than in the F-RM treatment. It should be noted, however, that the disclosure rate was more than 25% in the first half of the experiment, and as shown in Appendix Fig. C.4, the average payoff was then significantly larger at the 10% level for the non-disclosers than for the disclosers in the C-RM treatment.
[26] This tendency was observed throughout the entire experiment in the F-Sorting treatment. By contrast, in the C-Sorting treatment, this tendency was strong for the second half of the experiment (Appendix Fig. C.4).
[27] The disclosers' average payoffs are not significantly different between the two sorting treatments because although disclosers in the C-Sorting treatment achieved stronger cooperative relationships with their peers than those in the F-Sorting treatment, the former needed to incur a cost for their disclosure.
[28] Sorting increases the likelihood to be matched with someone who acts according to the conditional cooperation strategy, i.e., conditional cooperator or mimicker. The importance of the matching probability can be checked further using the data of the two random-matching treatments. As shown in Section D.1 of Appendix D, subjects' average disclosure rates were significantly positively correlated with their beliefs regarding the peers' disclosure rates. The subjects' conditional disclosure behaviors resonate with the idea that subjects would positively respond to

Result 3: *(i) The average contributions were higher for disclosers than for non-disclosers in all four treatments. (ii) The average payoffs were also higher for disclosers than for non-disclosers in the C-Sorting and F-Sorting treatments, but conversely, the payoffs were higher for non-disclosers than for disclosers in the C-RM and F-RM treatments.*[29]

Fig. 2: *Subjects' Average Contributions and Payoffs by Disclosure Decision*



Notes: Each bar was calculated in the following two steps. First, the session-average contributions (payoffs) were calculated by disclosure decision. Next, they were averaged by treatment. Each two-sided *p*-value in panel (a) (panel (b)) is a Wald test result for the null that the average contributions (payoffs) are the same between disclosers and non-disclosers in a given treatment. The Wald test was performed based on the estimation results of subject random effect linear regressions with standard errors clustered by session. The average contributions and payoffs in the Baseline and Mandatory treatments in the figure were also calculated using data from periods 2 to 20. *, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

As explained in Section 2, the Mandatory treatment was constructed for a comparison with the C-Sorting treatment. Fig. 2(a) indicates that the average contribution of disclosers in the C-Sorting treatment was almost the same as the average contribution in the Mandatory treatment, but that the average contribution of non-disclosers was much less. This suggests that people do

peers' signals of future contributions, as have been observed in other contexts, including direct punishment (Kamei, 2014), third party punishment (Kamei, forthcoming), and costly gossiping (Kamei and Putterman, 2018).

[29] Readers may wonder whether the treatment differences seen in Results 1 and 2 can go through in the same way if subjects are given an opportunity to repeat the supergame (finitely repeated public goods game). Kamei and Putterman (2017) found that subjects would (would not) learn to cooperate from supergame to supergame when cooperating is (is not) materially beneficial based on the partners' reciprocal responses. Result 3(ii) implies that the treatment differences between the two matching protocols may persist, or even widen, if subjects gain experiences.

cooperate if they are all forced to disclose, thus indicating that the mandatory disclosure system functions more effectively than the voluntary disclosure system since it eliminates the option to hide past behaviors.

5. Subjects' Disclosure Decisions and Action Choices

This section presents a detailed analysis of the subjects' behaviors.

5.1. Subjects' Disclosure Decisions and Reputation States

As discussed in Section 4.1 and Fig. 1(d), a non-negligible fraction of subjects did not disclose their reputation states (denoted as $S_{i,t}$, hereafter). This was the case even with sorting, despite Result 3(ii). In order to investigate this behavioral pattern in greater depth, the relationship between subjects' states and disclosure decisions was examined (Fig. 3). Three clear patterns emerged. First, a subject's likelihood to disclose $S_{i,t}$ is monotonically increasing in the size of $S_{i,t}$, whether disclosure is costly or free. This pattern, combined with Result 3(i), implies that a subject with a high (low) $S_{i,t}$ maintained a high (low) status over time, even though she could change her reputation status easily (recall that $S_{i,t} = c_{i,t-1}$).

Second, consistent with Hypothesis 2(a), sorting nevertheless raised subjects' disclosure rates for each category of $S_{i,t}$ except the highest. If $S_{i,t} \in [17, 20]$, the average disclosure rate was 94.5% even in the F-RM treatment.

Third, the presence of a positive disclosure cost discouraged subjects from disclosing $S_{i,t}$ in both the sorting and random-matching conditions, for each category of $S_{i,t}$ (see Result 1(iii) also).

Result 4: *(i) Subjects' disclosure rates were positively correlated with the size of own $S_{i,t}$. (ii) Even when disclosure was free, subjects with low $S_{i,t}$ were reluctant to disclose their states. (iii) The impact of sorting and of positive disclosure cost occurred regardless of the size of $S_{i,t}$.*[30]

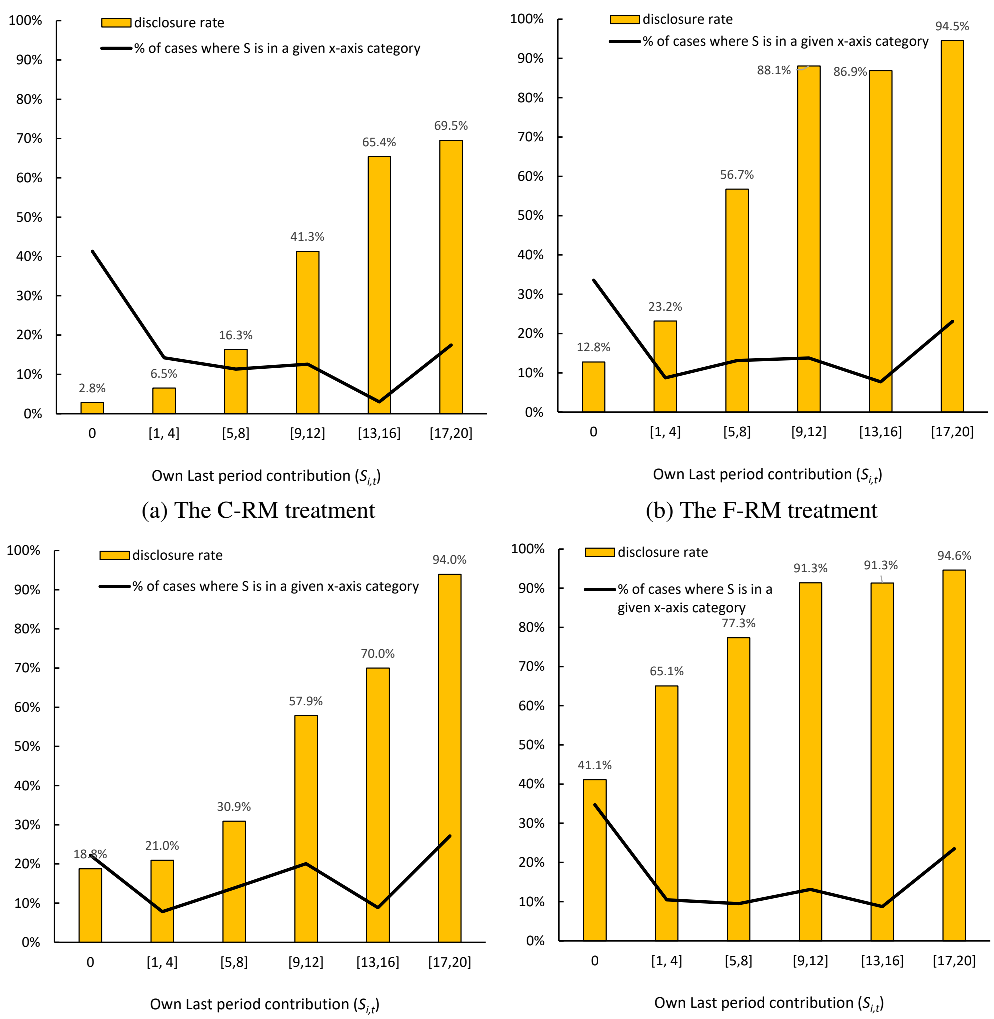
In summary, the weak rational cooperation under random matching (Result 2(i), Result 3(ii)), and the failure of information unraveling under sorting despite Result 3(ii), are due to some

---

[30] As shown in Appendix Fig. C.5, Result 4 holds irrespective of which half of the data (before period 11 or after period 10) are used.

subjects' hiding and uncooperative behaviors. But, why did they not attempt to build cooperative reputations using the information disclosure system? It is possible that non-disclosers mistakenly believed that their matched partners would not be skeptical about their intentions to hide (Jin *et al.*, 2015). An alternative explanation is bounded rationality and insufficient levels of reasoning (Benndorf *et al.*, 2015; Li and Schipper, 2018).[31]

Fig. 3: *Subjects' Disclosure Rates by the Size of their State*



(a) The C-RM treatment (b) The F-RM treatment

(c) The C-Sorting treatment (d) The F-Sorting treatment

Note: Data for periods 2 to 19 are used for this analysis, considering that the standard end-game effect was seen in the experiment.

[31] A regression analysis was conducted to explore whether it is possible to rationalize the subjects' behaviors with pecuniary considerations alone. It was found that in all treatments, (i) subjects' contribution amounts were positively correlated with their beliefs on peers' contribution behaviors and (ii) the beliefs were significantly positively correlated with the partners' states when these were observable. In the C-RM and F-RM treatments, in addition, subjects believed that non-disclosers' contribution behaviors would be the same as those of disclosers whose states $S_{i,t}$ were zero. The estimation results suggest that some subjects did not disclose even though disclosing the states was a materially beneficial action. The details are included in Section D.2 in Appendix D.

5.2. Cooperation Types and their Attempts to Strategically Build Reputation

The following considers how subjects' disclosure and contribution behaviors differ by cooperation type. As detailed below, the analysis reveals that (a) some selfish subjects strategically contributed positive amounts for future disclosure in line with the idea that they were mimicking the behaviors of conditional cooperators, but (b) not all conditional cooperators behaved cooperatively.

The conditional contribution schedule (Fischbacher *et al.*, 2001) elicited from each subject can be used for this analysis,[32] while focusing on two types: "conditional cooperators" and "free riders."[33] Similar to Fischbacher *et al.*, those whose own contributions were significantly positively correlated with the others' average contributions at least at the 5% level (according to spearman's ρ) are defined as the conditional cooperators. Those whose own contributions are always zero in the task are defined as free riders.[34] The percentages of conditional cooperators and free riders who disclosed their states in the C-RM (F-RM) treatment are on average 25.8% (55.3%) and 21.1% (50.3%), respectively. These two percentages in the C-Sorting (F-Sorting) treatment are 46.7% (70.6%) and 66.4% (65.8%), respectively. Thus, a non-negligible fraction of free riders did disclose their states, whereas not every conditional cooperator disclosed their state.

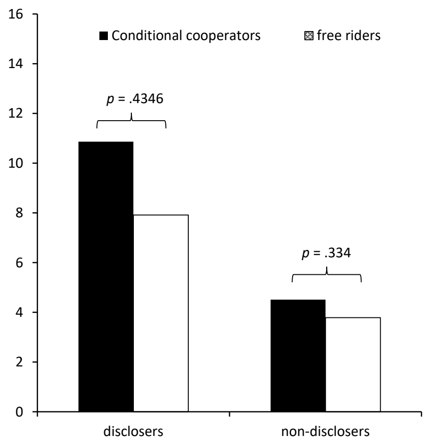
The fit of the Fischbacher *et al.*'s (2001) method can be checked by examining subjects' contribution amounts in period 20. A subject would contribute nothing in period 20 if she is purely selfish. The data show that free riders contributed much less in period 20 than conditional cooperators did, whether they disclosed the states or not (each panel (b) of Fig. 4). This implies that the classified types are good indicators to measure subjects' contribution behaviors.

The data by the type reveal that not only conditional cooperators but also free riders

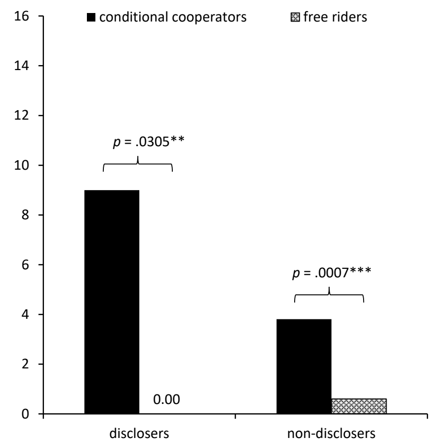
---

[32] The average conditional contribution schedule exhibits that of a conditional cooperator (Appendix Fig. C.3).
[33] Fischbacher *et al.* (2001) found a type called the "hump-shared contributor." Some subjects in this study also exhibited this pattern. However, only around 8.3% of subjects were classified as hump-shaped.
[34] 58.3% (63.6%) and 12.5% (20.5%) of the subjects were classified as conditional cooperators and free riders, respectively, in the C-RM (F-RM) treatment. 50.0% (59.1%) and 18.2% (18.2%) of subjects were classified as conditional cooperators and free riders, respectively, in the C-Sorting (F-Sorting) treatment.

Fig. 4: *Average Contribution Amounts by Disclosure Decision*

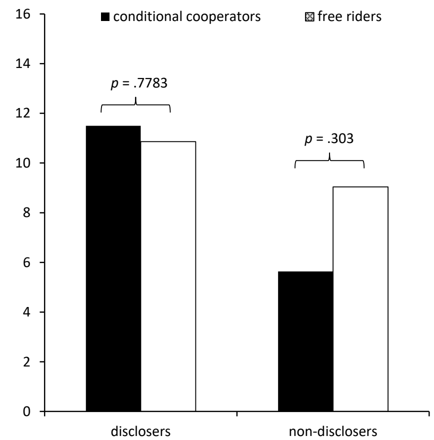


(a) Average before period 20

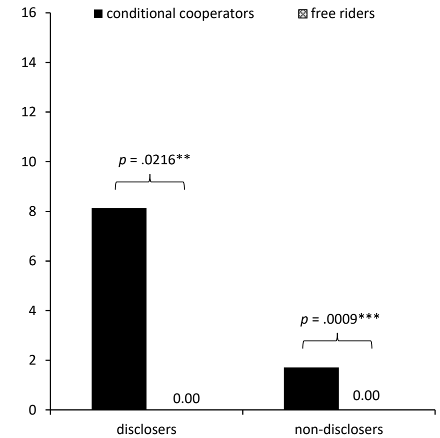


(b) Period 20 (final period)

(1) The C-RM treatment

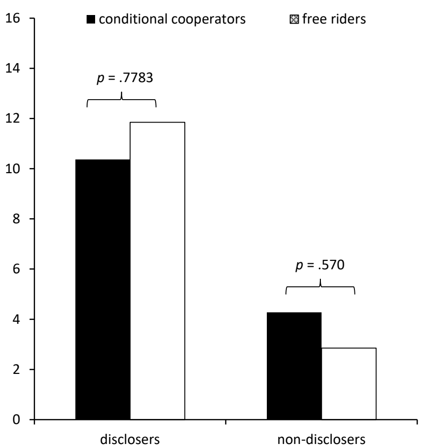


(a) Average before period 20

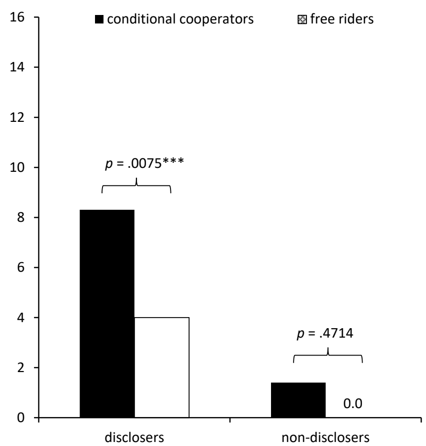


(b) Period 20 (final period)

(2) The F-RM treatment

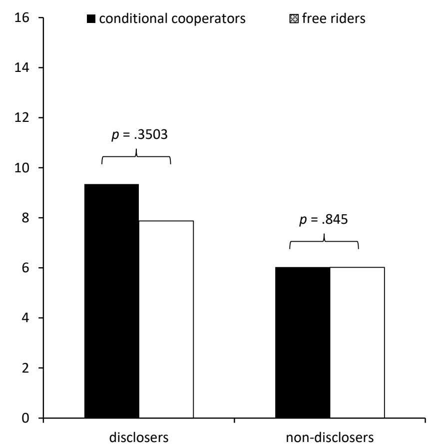


(a) Average before period 20

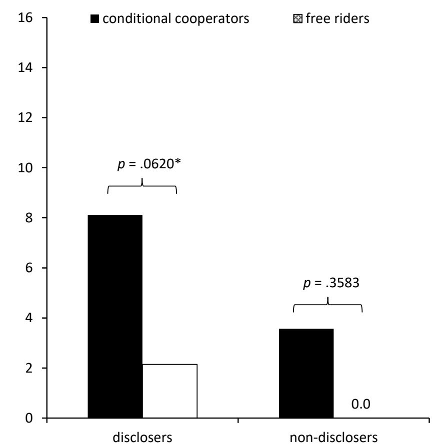


(b) Period 20 (final period)

(3) The C-Sorting treatment



(a) Average before period 20



(b) Period 20 (final period)

(4) The F-Sorting treatment

Notes: Each panel (a) reports average contribution amounts across all periods but periods 1 and 20 by disclosure decision. See Appendix Table C.4 for the calculations of the two-sided p-values reported in this figure. *, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

contributed similarly large amounts in each treatment before period 20 (each panel (a) of Fig. 4). This suggests that, whether the attempts were successful or not, some free riders mimicked the behavior of a conditional cooperator by contributing large amounts and disclosing them in the following periods.

Result 5: *(i) Free riders contributed very little in period 20, unlike conditional cooperators. (ii) However, the free riders contributed almost the same amounts as the conditional cooperators in periods before period 20.*

One may wonder why some conditional cooperators did not disclose their states. One possibility is that they were pessimistic about their peers' behaviors. As shown in Appendix Fig. D.2, conditional cooperators who less frequently disclosed their states formed lower beliefs about their matched peers' contribution behaviors, compared with the same types who frequently disclosed.[40] However, this is suggestive evidence only because the subjects' average beliefs do not differ significantly by the frequency of disclosure (also see Appendix Table D.2).

In the C-RM and F-RM treatments, disclosers were not assured to be matched with another discloser. Both conditional cooperators and free riders who disclosed their states contributed larger amounts when matched with disclosers rather than non-disclosers, but they still contributed large amounts even toward non-disclosers (Appendix Fig. C.6). This can be explained by the unmasked subjects' attempts to maintain high reputation scores ($S_{i,t+1}$ is solely dependent on i's contribution amount in period t). Interestingly, both cooperation types contributed larger amounts towards disclosers, even when they hid their states (each panel (b) of Fig. C.6). If we accept that free riders are strategically minded, this behavior by masked free riders may have been driven by an attempt to not discourage disclosers' willingness to contribute for future exploitative purposes.

5.3. Performance Differences between the C-Sorting and F-Sorting Treatments

Sorting had a positive effect on efficiency only when disclosure was costly (Fig. 1, Table 2,

[35] It should be noted that the direction of causality may be the opposite: subjects' beliefs on the disclosers' strong contribution behaviors may have been formed by their experiences.

Result 2(ii)). This subsection studies the difference between the two sorting treatments in details.

Appendix Fig. C.7 reports the subject-by-subject average contribution amounts when they chose to disclose the states. This shows that most subjects frequently switched back and forth between disclosing and not disclosing. However, more subjects selected to disclose frequently in the F-Sorting than in the C-Sorting treatment. Appendix Fig. C.8 reports the history of disclosure decisions, subject by subject. Unlike in the C-Sorting treatment, most subjects in the F-Sorting treatment switched back and forth between disclosing and hiding with different cycles.[36] The percentages of subjects who disclosed more than or equal to ten times (except period 20) – "frequent disclosers" hereafter – are 52.3% and 77.3% in the C-Sorting and F-Sorting treatments, respectively.[37] Fig. C.7 also indicates that a larger percentage of the frequent disclosers contributed small amounts in the F-Sorting than in the C-Sorting treatment.[38] These findings suggest that the presence of a positive disclosure cost effectively discouraged those who had intentions to exploit high contributors from disclosing their states. This implies that the positive effect of voluntary disclosure through better matching may not be large if disclosure is free.[39]

Result 6: *A significantly larger percentage of disclosers attempted to exploit high contributors in the F-Sorting than in the C-Sorting treatment.*

6. Conclusions

This study demonstrated that given an option to disclose the past, a non-negligible fraction of individuals do disclose their past action choices and that voluntary information disclosure helps

[36] Disclosers in period $t - 1$ in the C-Sorting treatment were more likely to stay as a discloser in period t, compared with those in the F-Sorting treatment (Remark in Appendix Fig. C.8). Non-disclosers in the C-Sorting treatment did not switch to being a discloser in the following period when they contributed large amounts (Appendix Table C.3).
[37] The two percentages are significantly different (two-sample test of proportions, two-sided p-value = .0141).
[38] The percentages of the frequent disclosers who contributed less than 10 ECUs are 11.4% and 43.2% in the C-Sorting and F-Sorting treatments, respectively. The two percentages are significantly different (two-sided p = .0107).
[39] As discussed in Section 2.2, the C-Sorting treatment has two components (A) disclosed information and (B) the cost to be matched with like-minded others. Component (B) is absent in the F-Sorting treatment. It could be considered that if the presence of a positive disclosure cost is the most important factor for the superior performance of the C-Sorting treatment, then the disclosers' benefits may remain high even if the element of disclosed information is eliminated from the C-Sorting treatment. The additional data from the C-Sorting-N treatment revealed that having a positive disclosure cost in fact generates an effect; however, the presence of disclosed information is crucial for performance in the C-Sorting treatment; see Appendix E for details.

strengthen cooperation. However, individuals were significantly deterred from disclosing if disclosure did not ensure being matched with another discloser, or if it involved a cost, even a small cost. The latter finding is surprising, considering the small size of the disclosure cost involved – only 1 ECU (2.5 pence). This may mean that there is a discontinuity in people's disclosure behaviors between zero and positive costs.

Nevertheless, costly information disclosure did support the evolution of cooperation when disclosers were assured to be matched with another discloser. This result supports past empirical research that emphasizes the role of sorting in sequential-move interactions (e.g., Tadelis and Zettelmeyer, 2015). This being said, the efficiency of the system as a whole under costly disclosure may be worse than that under mandatory disclosure, because (a) a sizable fraction of subjects conceal their states and behave uncooperatively and (b) disclosers need to incur a private cost. Aspect (a) may be caused by subjects' limited cognitive ability. For example, Li and Schipper (2018), in the context of a sequential-move sender-responder game, discussed that high levels of reasoning and cognitive abilities are required for first-movers in the persuasion game to recognize the benefits of quality disclosure. This argument also reinforces this study's conclusion that mandatory disclosure is superior, since it decreases subjects' cognitive loads. What players need to evaluate in the mandatory system is only the peers' responses to the value of their own states. Then, what may encourage those with less cognitive loads to voluntarily disclose their past behaviors and behave cooperatively? An exploration of mechanisms to further improve the costly information disclosure system remains an interesting avenue for future research.

REFERENCES

Andreoni, J. "Why free ride? Strategies and learning in public goods experiments." *Journal of Public Economics*, Vol. 37 (1988), pp. 291-304.

Benndorf, V., Kübler, D., and Normann H.-T. "Privacy concerns, voluntary disclosure of information, and unraveling: An experiment." *European Economic Review*, Vol. 75 (2015), pp. 43-59.

Bolton G., Katok, E., and Ockenfels, A. "How Effective are Online Reputation Mechanisms? An Experimental Study." *Management Science*, Vol. 50 (2004), pp. 1587-1602.

Bolton G., Katok, E., and Ockenfels, A. “Cooperation among strangers with limited information about reputation.” *Journal of Public Economics*, Vol. 89 (2005), pp. 1457-1468.

Camera, G., and Casari, M. “Cooperation among Strangers under the Shadow of the Future.” *American Economic Review*, Vol. 99 (2009), pp. 979-1005.

Camera, G., and Casari, M. “Monitoring institutions in indefinitely repeated games.” *Experimental Economics*, Vol. 21 (2018), pp. 673-691.

Chaudhuri, A. “Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature.” *Experimental Economics*, Vol. 14 (2011), pp. 47-83.

Dranove, D., and Jin, G.Z. “Quality Disclosure and Certification: Theory and Practice.” *Journal of Economic Literature*, Vol. 48 (2010), pp. 935-963.

Duffy, J., Xie, H., and Lee, Y.-J. “Social norms, information, and trust among strangers: theory and evidence.” *Economic Theory*, Vol. 52 (2013), pp. 669-708.

Engelmann, D., and Fischbacher, U. “Indirect Reciprocity and Strategic Reputation Building in an Experimental Helping Game.” *Games and Economic Behavior*, Vol. 67 (2009), pp. 399-407.

Fischbacher, U., and Gächter, S. “Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments.” *American Economic Review*, Vol. 100 (2010), pp. 541-56.

Fischbacher, U., Gächter, S., and Fehr, E. “Are people conditionally cooperative? Evidence from a public goods experiment.” *Economics Letters*, Vol. 71 (2001), pp. 397-404.

Forsythe, R., Isaac, M., and Palfrey, T. “Theories and Tests of “Blind Bidding” in Sealed-bid Auctions.” *Rand Journal of Economics*, Vol. 20 (1989), pp. 214-238.

Gächter, S., and Renner, E. “The effects of (incentivized) belief elicitation in public goods experiments.” *Experimental Economics*, Vol. 13 (2010), pp. 364-377.

Hagenbach, J., and Perez-Richet, E. “Communication with Evidence in the Lab.” working paper, 2018.

Hatton, C. “China 'social credit': Beijing sets up huge system.” *BBC News*, October 26, 2015.

Honhon, D., and Hyndman, K., 2019. Flexibility and Reputation in Repeated Prisoner's Dilemma Games. Working paper.

Jin, G., Luca, M., and Martin, D. “Is No News (Perceived as) Bad News? An Experimental Investigation of Information Disclosure.” Harvard Business School Working Paper, No. 15-078, 2015.

Kamei, K. “Conditional Punishment.” *Economics Letters*, Vol. 124 (2014), pp. 199-202.

Kamei, K. “Endogenous Reputation Formation under the Shadow of the Future.” *Journal of Economic Behavior & Organization*, Vol. 142 (2017), pp. 189-204.

Kamei, K. “The power of joint decision-making in a finitely-repeated dilemma.” *Oxford Economic Papers*, Vol. 71 (2019), pp. 600-622.

Kamei, K. “Group Size Effect and Over-Punishment in the Case of Third Party Enforcement of Social Norms.” *Journal of Economic Behavior & Organization* (forthcoming).

Kamei, K., and Putterman, L. “Play it Again: Partner Choice, Reputation Building and Learning from Finitely-Repeated Dilemma Games.” *Economic Journal*, Vol. 127 (2017), pp. 1069-1095.

Kamei, K., and Putterman, P. “Reputation Transmission without Benefit to the Reporter: a Behavioral Underpinning of Markets in Experimental Focus.” *Economic Inquiry*, Vol. 56 (2018), pp. 158-172.

Kandori, M. “Social Norms and Community Enforcement.” *Review of Economic Studies*, Vol. 59 (1992), pp. 63-80.

Kreps, D., Milgrom, P., Roberts, J., and Wilson, R. “Rational Cooperation in the Finitely Repeated Prisoners’ Dilemma.” *Journal of Economic Theory*, Vol. 27 (1982), pp. 245-252.

Ledyard, J. “Public goods: A survey of experimental research.” In J. H. Kagel and A.E. Roth, eds., *The Handbook of Experimental Economics* 111-194, Princeton University Press, 1995.

Li, Y., and Schipper, B. “Strategic Reasoning in Persuasion Games: An Experiment.” working paper, 2018.

Page, T., Putterman, L., and Unel, B. “Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency.” *Economic Journal*, Vol. 115 (2005), pp. 1032-1053.

Shampanier, K., Mazar, N., and Ariely, D. “Zero as a Special Price: The True Value of Free Products.” *Marketing Science*, Vol. 26 (2007), pp. 742-757.

Stahl, D. “An experimental test of the efficacy of a simple reputation mechanism to solve social dilemmas.” *Journal of Economic Behavior & Organization*, Vol. 94 (2013), pp. 116-124.

Tadelis, S., and Zettelmeyer, F. “Information Disclosure as a Matching Mechanism: Theory and Evidence from a Field Experiment.” *American Economic Review*, Vol. 105 (2015), pp. 886-905.

Online Supplementary Appendix for: Kamei, Kenju

"Voluntary Disclosure of Information and Cooperation in

Simultaneous-Move Economic Interactions"[1]

Table of Contents

[1] Correspondence: Kenju Kamei, Department of Economics and Finance, Durham University, Mill Hill Lane, Durham, DH1 3LB, United Kingdom. Email: kenju.kamei[at]gmail.com, kenju.kamei[at]durham.ac.uk.

Appendix A: Some Experimental Procedure and Instructions

The experiment except the instructions was programmed using the z-Tree software (Fischbacher, 2007). The instructions were neutrally framed. Eligible subjects were sent solicitation messages via ORSEE developed by Greiner (2015); and subjects voluntarily registered for and participated in the experiment.

References:

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

Greiner, B. "Subject pool recruitment procedures: organizing experiments with ORSEE." *Journal of Economic Science Association*, Vol. 1 (2015), pp. 114-125.

This part of the Appendix includes instructions for the C-Sorting treatment as an example.

[The following instructions were read aloud at the onset of the experiment:]

Instructions

You are now taking part in a decision-making experiment. Depending on your decisions and the decisions of other participants, you will be able to earn money in addition to the £3 guaranteed for your participation. Please read the following instructions carefully.

During the experiment you are not allowed to communicate with other participants. If you have questions, raise your hand. One of us will come to answer your question.

During the experiment your earnings will be calculated in points. At the end of the experiment your points will be converted to U.K. pounds at the following rate:

40 points = £1

(or each point will be exchanged for 2.5 pence of real money). At the end of the experiment your total earnings (including the **£3** participation fee) will be paid out to you in cash. Your payment will be rounded to the nearest 10 pence (e.g., £15.30 if it is £15.33; and £15.40 if it is £15.37).

This experiment consists of two parts. We will first explain the detail of Part 1. We will distribute the instruction for Part 2 once Part 1 is over.

PART 1

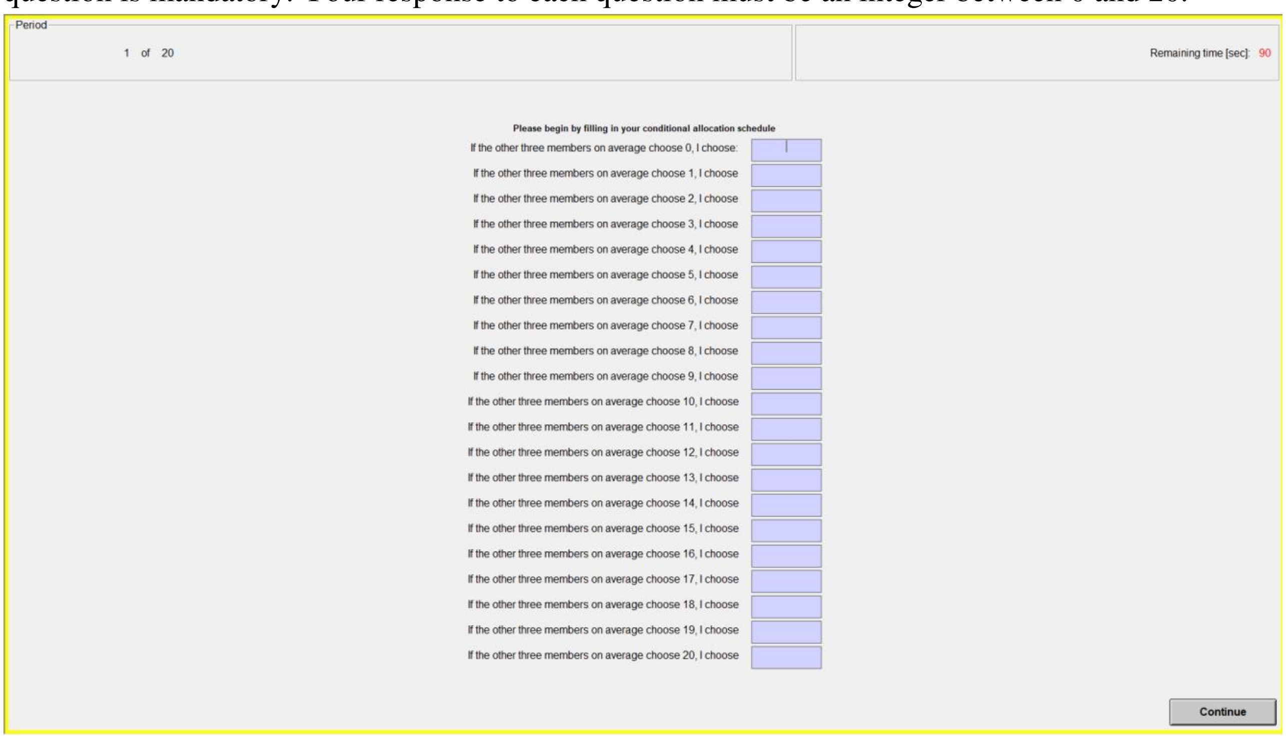
At the beginning of Part 1, you are randomly assigned to a group of four and interact with each other. In this part, you and your three group members are each given **an endowment of 20 points** and simultaneously make two kinds of allocation decisions. There are two accounts to allocate: **private account and group account**. Specifically, you are asked how many points you want to allocate to the group account. The remaining points (that is, 20 minus your allocation to the group account) will be automatically allocated to your private account. Your earnings in this part depend on **(a) the number of points in your private account and (b) total allocation amounts to your group account**.

You are asked to make the following two decisions in order. We will first explain these two decisions. We will then explain how your earnings are determined.

The first decision: *conditional allocation decisions*

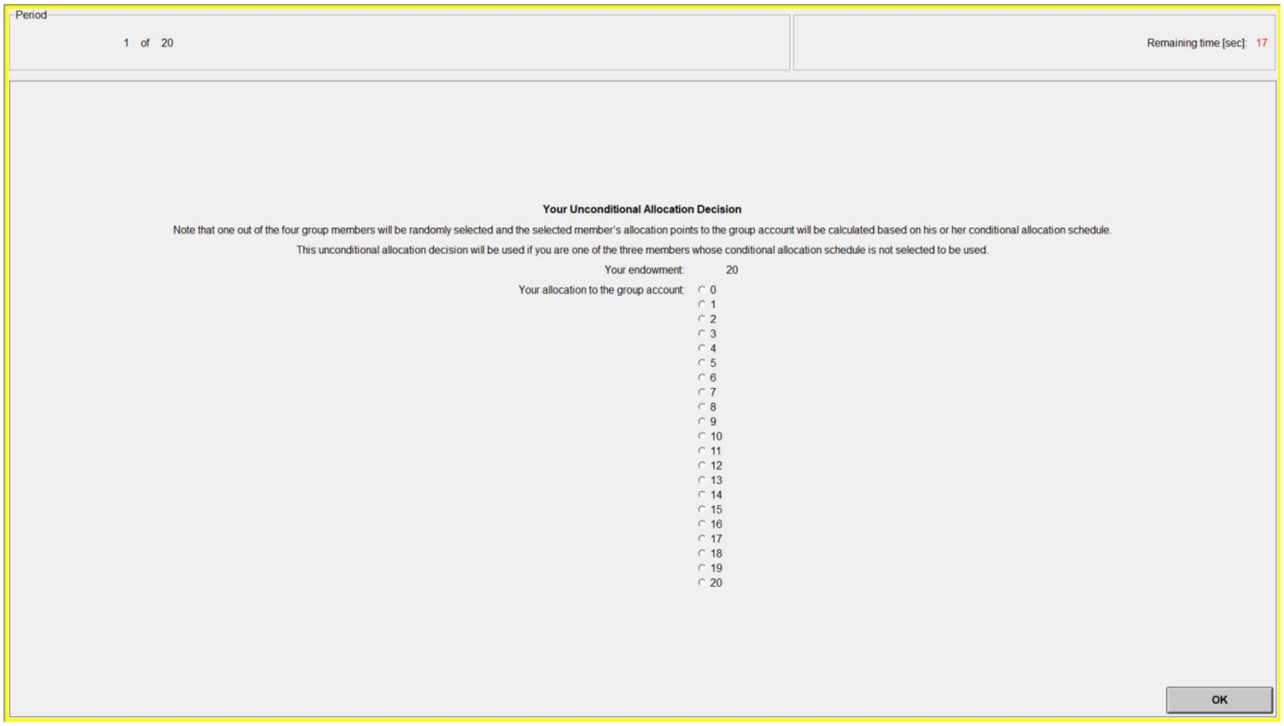
You are asked how many points you want to allocate to your group account, contingent on average allocations of the three other members in your group. Specifically, you are asked to make 21 conditional decisions by completing the form shown below:

For instance, in the top box, you will input how many points you want to allocate to your group account if the other three members on average allocate 0 points to your group account. Likewise, in the bottom box, you will answer how many points you want to allocate to your group account if the other three members on average allocate 20 points to your group account. Answering each question is mandatory. Your response to each question must be an integer between 0 and 20.



The second decision: *unconditional allocation decisions*

Once all participants complete the conditional allocation schedules, you will be subsequently asked to make unconditional allocation decisions to their group accounts.



How to calculate your earnings:

First, both of your conditional and unconditional allocation decisions may affect your earnings. Once all individuals in your group make their unconditional allocation decisions, one out of the four group members is randomly selected and the selected member's allocation points to the group account will be calculated based on his or her conditional allocation schedule. As for the other three participants in the group, their unconditional allocation decisions will be used. Note that in calculating the selected member's allocation points, the other three members' average allocation amount is rounded to the nearest integer (e.g., 4 if it is 3.5; 10 if it is 10.4).

For example, suppose that you were randomly chosen as the one whose conditional allocation schedule will be used. Also, suppose that the three other members in your group unconditionally allocated 5 points, 11 points and 8 points, respectively, to the group account. Then, the average allocation amount of the three members is calculated as 8 points (= (5 + 11 + 8)/3). Suppose that you indicated that you would allocate 10 points if the others on average allocated 8 points to your group account. Then, your allocation amount in Part 1 would be 10 points.

Your earnings in this part are calculated as in the following formula:

(sum of points in your private account) + 0.4 × (sum of points allocated by you and your group members to the group account)

In other words, your earnings from your private account **are equal to the number of points you allocated to the private account** (20 minus your allocation to your group account). The points you allocate to your private account do not affect the earnings of your group members.

By contrast, your earnings from the group account equal the **sum** of points allocated to the group account by you and your three group members multiplied **by 0.4**. In other words, **when you allocate 1 point to the group account, you and your three peers each get 0.4 (= 1 × .4) points as earnings**. Thus, the total earnings in this case are 1.6 points, which is greater than 1 point. Note that you also obtain earnings of 0.4 points for each point your group member allocates to your group account.

You will be informed of the interaction outcomes and earnings of Part 1 after Part 2 is over.

If you have any questions, please raise your hand. When all questions are answered, we will move on to comprehension questions.

Comprehension questions

Please answer the following questions. Raise your hand if you need help.

1. Suppose that all four members in your group allocate 0 points to the group account. How much does each member earn? _______________

2. Suppose that all four members in your group allocate 20 points to the group account. How much does each member earn? _______________

3. Suppose that the other three members in your group in total allocate 30 points to the group. Answer the following:

a) How much do you earn if you allocate 0 points to the group account? _______________

b) How much do you earn if you allocate 10 points to the group account? _______________

c) How much do you earn if you allocate 20 points to the group account? _______________

Any questions? When all questions are answered, we will move on to Part 1.

[Once everyone finished answering the comprehension questions in Part 1 and the experimenter explained the answers, Part 1 began. Once Part 1 was over, the following instructions were distributed and were read aloud:]

PART 2

In this part, **you will be paired and interact with another participant in each period**. The number of interactions in Part 2 is 20. At the beginning of each period, you are given an endowment of 20 points and make a decision using the endowment. You will be paid based on your accumulated earnings at the end of the experiment (the conversion rate is 40 points = £1 as in Part 1).

Each period after period 1 consists of two stages. (Period 1 consists of only one stage.) We will first explain the nature of your interactions in each period. We will then explain how your partner is assigned to you.

Your interactions in each period:

You and your assigned counterpart will be each given an **endowment of 20 points** in every period. You and your partner then simultaneously decide how to use the endowment. There are two accounts for this purpose: **private account and joint account**. You will be asked how many points you want to allocate to your joint account. The remaining points will automatically be allocated to your private account.

Your earnings in a given period depend on **(a) the number of points in your private account and (b) the total number of points in the joint account**.

How to calculate your earnings:

Your earnings from your private account **are equal to the number of points allocated to the private account**. That is, for example, if there are 5 points in your private account, you get 5 points as earnings. The points you allocate to your private account <u>do not</u> affect the earnings of your counterpart.

Your earnings from the joint account equal the **sum** of points allocated to the joint account by you and your counterpart multiplied **by 0.8**. In other words, **when you allocate 1 point to the joint account, you and your partner each get 0.8 (= 1 × .8) points as earnings**. For example, suppose that you decide to allocate 5 points to the joint account. Also suppose that your counterpart decides to allocate 9 points to the joint account. In this case, the sum of points in the joint account is 14 (= 5 + 9). The earnings of you and your counterpart from the joint account are equal to 11.2 (= 14 × 0.8) points.

In summary, your earnings in a given period are calculated with the following formula:

(sum of points in your private account) + 0.8 × (sum of points allocated by you and your counterpart to the joint account)

Note that you get 1 point as earnings when you allocate 1 point to your private account. If you instead allocate 1 point to your joint account, your earnings from your allocation is 0.8 × 1 = 0.8 points. However, by allocating 1 point to the joint account, the earnings of your counterpart also increase by 0.8 points. Therefore, the total pair earnings are 0.8 × 2 = 1.6 points, which is greater than 1 point. Note that you also obtain earnings from points allocated to the joint account by your counterpart. You obtain 0.8 × 1 = 0.8 points for each point your pair partner allocates to your joint account.

How is your partner assigned to you in each period?

At the onset of each period, all participants simultaneously **decide whether to disclose how much they allocated to their joint account in the last period by paying 1 point**. If you decide to disclose it, 1 point will be deducted from your given period's earnings.

Each person that chose to disclose will be randomly matched **with another person that decided to disclose**. **His or her allocation amount to the joint account in the last period is then informed to his or her partner before the partner makes an allocation decision**. If the number of those who chose to disclose their last-period allocation amounts is an odd number, then one of them will be randomly matched with a person that chose not to disclose his or her last-period allocation. Except the one person, each person that decided to disclose is assured that they will be matched with another person that chose to disclose.

Likewise, each person that decided not to disclose his or her allocation amount to the joint account in the last period **will be randomly matched with another person that likewise chose not to disclose**. If the number of those who did not disclose their last-period allocation amount is an odd number, then one of them will be randomly matched with a person that chose to disclose his or her last-period allocation amount. Except the one person, each person that decided not to disclose is assured that they will be randomly matched with another person that decided not to disclose.

There is no disclosing decision for you to make in period 1 since there is no previous period in period 1. Each participant is randomly matched with another participant in that period in the today's experiment. In period 2, you decide whether to disclose your first period allocation amount which costs you 1 point; and your matching is determined based on your disclosure decision and the computer's random choices as explained above.

Once this matching process in a given period is completed, you will move on to the allocation decisions as already described above.

Summary:

In period 1, you will be randomly matched with one of other participants, are given an endowment of 20 points, and make an allocation decision between your private account and the joint account. Your allocation decision and your counterpart's allocation decision determine your earnings in period 1.

In each period after period 1, you have two stages.

Stage 1: All participants simultaneously decide **whether to disclose** how much they allocated to their joint accounts in the last period **by paying 1 point**. Each participant that decided to disclose his or her last-period allocation will be randomly matched with another that also decided to disclose his or her last-period allocation amount. Likewise, each participant that decided not to disclose his or her last-period allocation amount is randomly matched with another participant that also decided not to disclose his or her last-period allocation amount.

Stage 2: Both you and your counterpart simultaneously make allocation decisions using the assigned endowment of 20 points. Your earnings in a given period are determined by your and your counterpart's allocation decisions.

You are privately paid based on your accumulated earnings immediately after the experiment is over.

If you have any questions, please raise your hand. When all questions are answered, we will move on to comprehension questions.

Comprehension questions

Please answer the following questions. Raise your hand if you need help.

1. How many periods do you have in Part 2? ________________

2. How many points are deducted from your earnings at the end of a period if you decide to disclose your last-period allocation decision to your partner in that period? ________________
How many points are deducted from your earnings at the end of a period if you decide not to disclose your last-period allocation decision? ________________

3. Suppose that both you and your counterpart allocate 0 points to the joint account in a given period.

a) How much do you earn? _______________

b) How much does your counterpart earn? _______________

4. Suppose that both you and your counterpart allocate 20 points to the joint account in a given period.

a) How much do you earn? _______________

b) How much does your counterpart earn? _______________

5. Suppose that your counterpart allocates 10 points to the joint account in a given period. Answer the following:

a) How much do you earn if you allocate 0 points to the joint account? _______________

b) How much do you earn if you allocate 10 points to the joint account? _______________

c) How much do you earn if you allocate 20 points to the joint account? _______________

Any questions?

[Once everyone finished answering the comprehension questions and the experimenter explained the answers, Part 2 started.]

Appendix B: Theoretical Calculations

Unlike Hypothesis 1 of the paper, cooperation could evolve with voluntary information disclosure. In the framework used in this study, a typical example of a non-selfish discriminating strategy is the 'conditional cooperation strategy' (e.g., Fischbacher *et al.*, 2001).[2] The prevalence of conditional cooperators (i.e., players who follow the conditional cooperation strategy) is well-documented. In the theoretical analysis, for simplicity, assume that a conditional cooperator m contributes $x \cdot S_{j,t}$, where $x \in (0,1]$, in period t if her matched partner j's state (last-period contribution) is $S_{j,t} = c_{j,t-1}$ and it is observable; m contributes zero if the state is not observable.[3] The steady decline of contributions in a finitely repeated dilemma game with no institutions can then be interpreted as a phenomenon that emerges when conditional cooperators are discouraged from cooperating after witnessing selfish types free ride (e.g., Fischbacher and Gächter, 2010).

The following discusses the possibility of rational cooperation by examining the optimal behavior of a strategic free rider type i, assuming that there exist players who choose to disclose and then act according to the conditional cooperation strategy discussed above, in addition to the other free rider types. $p_{cc} \times 100\%$ and $p_{FR} \times 100\%$ are used to refer to the percentages of the conditional cooperators and the free riders, respectively ($p_{cc} + p_{FR} = 1$). Also, assume that $p_{m,k} \times 100\%$ of the free rider types mimic the behavior of a conditional cooperator ($(1 - p_{m,k}) \times 100\%$ of the free rider types contribute zero unconditionally) in treatment k. Hereafter, $p_{m,k}^{RM}$, $p_{m,k}^{S}$ and $p_{m,k}^{M}$ are used to refer to $p_{m,k}$ under random matching, sorting, and mandatory disclosure, respectively. Further, assume that $x = 1$ (perfect conditional cooperator) for simplicity.[4] In this illustrative framework, we can show that it is materially beneficial for i to mimic the behavior of a conditional cooperator if there is a sufficiently large fraction of conditional cooperators and mimickers in the community where disclosers operate. The detail is provided in Sections B.1 and B.2. This prediction can be derived by setting up a Hamiltonian and applying the Maximum principle to it (e.g., Sethi and Thompson, 2006).

Proposition B.1: *If the percentage of persons who act according to the conditional cooperation strategy (i.e., conditional cooperators or mimickers) in the community where disclosers operate, denoted as "p", is sufficiently large that $p > 0.25$, then: in all periods before period 20, i contributes the full endowment and then discloses the state; in the end period (period 20), i*

---

[2] Such conditional behaviors can be rationalized, for example, by assuming that players are concerned about inequity (e.g., Fehr and Schmidt, 1999) or intention-based reciprocity (e.g., Rabin, 1993; Dufwenberg and Kirchsteiger, 2004).

[3] In sealed-bid auctions with sellers' voluntary disclosure of product quality, Forsythe *et al.* (1989) concluded that a sequential equilibrium in which buyers pessimistically believe that the worst scenario would happen for the blind bid explains subjects' behaviors the most accurately. The assumption that conditional cooperators contribute zero if they encounter non-disclosers is similar to the "assume-the-worst strategy" employed by the subjects in Forsythe *et al.* (1989) because the "worst" partner in the present study is someone who contributes nothing.

[4] The basic implication for a strategic free rider's behaviors does not change even if $x < 1$ is assumed. In this theoretical analysis, it is also assumed that each discloser (non-discloser) is randomly matched with another discloser (non-discloser) always under the sorting condition for simplicity.

discloses the state and then selects $c_{i,20} = 0$. If $p < 0.25$, the prediction is the same as Hypothesis 1 of the paper.

The treatment differences depend on assumptions. Assume that the free riders who do not mimic make disclosure decisions stochastically with a probability of y_k in treatment k. Lastly, suppose that i correctly anticipates her peers' behaviors and her beliefs on p_{cc}, $p_{m,k}$ and y_k are correct. Under these assumptions, p in Proposition B.1 is calculated as follows:

$$\text{(A)}\begin{cases} p = p_{cc} + p_{FR} \cdot p_{m,k}^{RM} \equiv p^{RM} \text{ for both the C-RM and F-RM treatments.} \\ p = \frac{p_{cc}+p_{FR}\cdot p_{m,k}^{S}}{p_{cc}+p_{FR}\cdot p_{m,k}^{S}+y_k\cdot p_{FR}\cdot(1-p_{m,k}^{S})} \equiv p^{S} \text{ for the C-Sorting and F-Sorting treatments.} \\ p = p_{cc} + p_{FR} \cdot p_{m,M}^{M} \equiv p^{M} \text{ for the Mandatory treatment.} \end{cases}$$

Here, we can reasonably assume that $p_{m,k}^{RM}$ is the same for the C-RM and F-RM treatments since the disclosure cost is too small to affect i's mimicking incentives (see the mathematical calculations summarized in Section B.1). Since past experimental studies suggest that endogenous regrouping or partner choice encourages selfish types to mimic cooperative types, it can also be assumed that $p_{m,k}^{S} > p_{m,k}^{RM}$ for a given disclosure cost.

Condition (A) and Proposition B.1 suggest three behavioral patterns. First, i's incentives to build a cooperative reputation would be stronger with than without sorting.[5] Second, the disclosure cost does not affect subjects' contribution behavior under random matching, as the cost is too small: just one ECU.[6] Third, mandatory disclosure would improve cooperation, compared with random matching, because prior research suggests that $p_{m,M}^{M} > p_{m,k}^{RM}$ (and thus $p^{M} > p^{RM}$). Kamei (2017), in the context of an indefinitely repeated prisoner's dilemma game, showed that: (i) given an option to hide IDs, a non-negligible fraction of individuals do not disclose their IDs and behave uncooperatively, even if mutual cooperation holds as an equilibrium outcome; and (ii) some subjects do not learn to disclose, even though disclosers continue to select defection against masked subjects. Such hiding and uncooperative behaviors might be even stronger in the present study due to the finitely repeated setup. Kamei (2017) also showed that eliminating the option to hide would increase the number of mimickers. These considerations can be summarized as in Hypothesis 2 of the paper. It is noted here that no clear prediction is possible for a comparison between the sorting treatments versus the Mandatory treatment. $p_{m,k}^{S}$ may be larger than $p_{m,M}^{M}$ considering that sorting assures a better matching among like-minded individuals. However, the opposite ($p_{m,M}^{M} > p_{m,k}^{S}$) may also be possible if eliminating the option to hide encourages those who would hide and then

---

[5] $p^{S} - p^{RM} = \frac{p_{cc}+p_{FR}\cdot p_{m,k}^{S}}{p_{cc}+p_{FR}\cdot p_{m,k}^{S}+y_k\cdot p_{FR}\cdot\left(1-p_{m,k}^{S}\right)} - \left(p_{cc} + p_{FR} \cdot p_{m,k}^{RM}\right) \geq \left(p_{cc} + p_{FR} \cdot p_{m,k}^{RM}\right)\left[\frac{1}{p_{cc}+p_{FR}\cdot p_{m,k}^{S}+y_k\cdot p_{FR}\cdot\left(1-p_{m,k}^{S}\right)} - 1\right] \geq 0.$

[6] A positive disclosure cost may raise the mimicking incentives under sorting if the cost deters some opportunistic types from invading the set of disclosers. As an anonymous referee pointed out, however, the impact of cost depends on assumptions regarding free riders' mimicking behaviors (a full characterization of equilibrium without imposing assumptions is desirable). A rigorous theoretical analysis is, however, beyond the scope of this paper.

behave uncooperatively in the sorting treatments to strategically build a reputation under mandatory disclosure.

The rest of Appendix B provides calculation results. It first explains a strategic selfish free rider *i*'s optimal control, assuming that all free riders other than *i* do not mimic the behavior of a conditional cooperator (Section B.1). It will then explain *i*'s optimal control when some other free riders also pretend to be a conditional cooperator (Section B.2).

B.1. When no free riders other than *i* are mimickers

Notation:

$p_{cc}\times 100$: the percentage of conditional cooperators (those who disclose last-period contribution amounts and then act according to the conditional cooperation strategy)

$p_{FR}\times 100$: the percentage of free riders other than *i* (those who contribute 0 unconditionally)

f: disclosure cost (= 1 in the C-RM and C-Sorting treatments; = 0 in the F-RM and F-Sorting treatments)

$\lambda_{i,t+1}$: the shadow price of a unit of the reputation state in period $t + 1$

Assumption: $y_k\times 100\%$ of free rider types disclose their states (although they contribute zero) in the treatment $k \in$ {C-RM, F-RM, C-Sorting, F-Sorting}.

* Note that having unmasked free riders would not affect calculation results under random matching, because of the assumption that conditional cooperators contribute zero towards disclosers whose states are zero, but it would decrease the value of information disclosure under sorting. The disclosure behavior suggested by Hypothesis 1 of the paper (standard theory prediction) means $y_{C\text{-}RM} = y_{C\text{-}Sorting} = 0$ and $y_{F\text{-}RM} = y_{F\text{-}Sorting} = 0.5$.

As summarized below, calculations indicate the following treatment differences:

- Provided that there is a sufficiently large percentage of conditional cooperators, the presence of a disclosure cost does *not* affect subjects' decisions to build cooperative reputations, because the disclosure cost is sufficiently small (1 ECU). Further, the threshold percentage of conditional cooperator, $\tilde{p}_{cc}$, above which *i* mimics the behavior of a conditional cooperator is also small. For example, $\tilde{p}_{cc}$= .25 in the C-RM and F-RM treatments.
- Suppose that there is a sufficiently large fraction of conditional cooperators. Free rider *i* then selects to disclose her state and always selects a contribution level of 20 until period 19 (i.e., the Bang-bang solution).
- Subjects' average contribution amounts are higher in the C-Sorting and F-Sorting than in the C-RM and F-RM treatments, respectively, because each discloser is assured to be matched with another discloser under the sorting condition.

- Subjects' contribution behaviors in the Mandatory treatment are the same as those in the C-RM and F-RM treatments.

B.1.1. *C-RM and F-RM treatments*

Assume that every subject has a reputation score of c_0 (initial expected reputation score) in period 1 since they are not given an option to disclose in that period. This assumption means that conditional cooperators contribute $S_1 = c_0$ (while free riders contribute zero) in period 1. A Hamiltonian function can then be set up as in (#0):

$$(\#0)\quad \begin{cases} H_{t,i} = 20 - c_{i,t} + r\cdot(c_{i,t} + [p_{cc}\cdot S_1 + (1 - p_{cc})\cdot 0]) + \lambda_{i,t+1}\cdot\Delta S_{i,1}, \text{ for } t = 1. \\ H_{t,i} = 20 - c_{i,t} + r\cdot(c_{i,t} + p_{cc}\cdot S_{i,t}\cdot 1_{\text{disclose},i,t}) - f\cdot 1_{\text{disclose},t} + \lambda_{i,t+1}\cdot\Delta S_{i,t}, \text{ for } t > 1. \end{cases}$$

This control problem for i is written as:

$$\max_{(c_{i,t})_{t\le 20}} \left\{J = \textstyle\sum_{t=1}^{20}[\pi(i,t) = H_{i,t} - \lambda_{i,t+1}\Delta S_{i,t}]\right\}, \text{ subject to:}$$

$$\Delta S_{i,t} = S_{i,t+1} - S_{i,t} = c_{i,t} - S_{i,t} \text{ for } t = 1, 2, \ldots, 20; S_{i,1} = S_1.$$

Under this setup, the adjoint equation can be written as below:

(#1)

- For $t = 2, \ldots, 20$, $\Delta\lambda_{i,t} = \lambda_{i,t+1} - \lambda_{i,t} = -\partial H_{t,i}/\partial S_{i,t} = -p_{cc}\cdot r\cdot 1_{\text{disclose},t} - \lambda_{i,t+1}\cdot(-1)$. In other words,

$$\lambda_{i,t} = p_{cc}\cdot r\cdot 1_{\text{disclose},t}.$$

- For $t = 21$, $\lambda_{i,21} = \partial\pi(i,21)/\partial S_{i,20} = 0$ (notice that there is no period 21 in the experiment).

The optimal control for $\{c_{i,t}\}_{t\le 20}$ is a bang-bang control, because $\partial H_{t,i}/\partial c_{i,t} = -1 + r + \lambda_{i,t+1}$ does not depend on $c_{i,t}$. Specifically, from the above adjoint equations, there are two cases:

(#2)

- When $t < 20$, $\partial H_{t,i}/\partial c_{i,t} = -1 + r + \lambda_{t+1} = -1 + r + p_{cc}\cdot r\cdot 1_{\text{disclose},t+1}$. This means that $c_{i,t} = 20$ if $r + p_{cc}\cdot r\cdot 1_{\text{disclose},t+1} > 1$. Since $r = .8$, this inequality reduces to: $p_{cc}\cdot 1_{\text{disclose},t+1} > .25$. By contrast, $c_{i,t} = 0$ if $-1 + r + p_{cc}\cdot r\cdot 1_{\text{disclose},t+1} < 0$, or $p_{cc}\cdot 1_{\text{disclose},t+1} < .25$.

- When $t = 20$, $\partial H_{t,i}/\partial c_{i,t} = -1 + r$, which is negative always. Thus, $c_{i,20} = 0$.

The optimal disclosure decision of subject i is dependent on $S_{i,t}$. Equation (#0) suggests that:

- i discloses (does not disclose) $S_{i,t}$ if $S_{i,t} > (<) f/(p_{cc}\cdot r) = 1.25/p_{cc}$; i randomly decides whether to disclose the state if $S_{i,t} = 1.25/p_{cc}$, in the C-RM treatment;

- i discloses (randomly decides whether to disclose) $S_{i,t}$ if $S_{i,t} \neq 0$ ($S_{i,t} = 0$) in the F-RM treatment.

These analyses suggest that in the F-RM treatment, subject i always discloses and contributes 20 so long as $p_{cc} > .25$ until period 19; and then i contributes zero in period 20. By contrast, if $p_{cc} < .25$, i always contributes zero and randomly decides whether to disclose in each period.

Because the disclosure cost is just one ECU, subjects' reputation building behaviors in the C-RM treatment would be the same as in the F-RM treatment. Because of the bang-bang solution, when $p_{cc} > .25$ (then $c_{i,t} = 20$), the condition of $S_{i,t} > 1.25/p_{cc}$ reduces to: $p_{cc} > 1.25/20 = .0625$. This condition automatically holds due to the assumption that $p_{cc} > .25$. Thus, in the C-RM treatment, if $p_{cc} > .25$, subject i always discloses and contributes 20 until period 19; and i disclose $S_{i,20}$ but contributes nothing in period 20.

If $p_{cc} < .25$, condition (#2) suggests that $c_{i,t} = 0$ for all t in the C-RM treatment. With this optimal control, it is beneficial for i to not disclose $S_{i,t}$ since $S_{i,t} > f/(p_{cc} \cdot r)$ does not hold.

Summary 1: *(a) The presence of a positive disclosure cost does not affect the free rider i's incentive to mimic the behavior of a conditional cooperator under random matching.*

(b) If there is such a large fraction of conditional cooperators that $p_{cc} > .25$ *in the population, i always strategically contributes 20 and then discloses the state until period 19 in the C-RM and F-RM treatments. In period 20, i discloses the state but selects* $c_{i,20} = 0$.

(c) If $p_{cc} < .25$*, i contributes 0 in all periods in the C-RM and F-RM treatments; i never discloses the state in the C-RM treatment, while i randomly decides whether to disclose the state in each period in the F-RM treatment.*

B.1.2. *C-Sorting and F-Sorting treatments*

In the two sorting treatments, disclosers (non-disclosers) are assured to be matched with disclosers (non-disclosers). Thus, the likelihood to interact with a conditional cooperator when i discloses her state is larger, compared with in the corresponding random matching environments. The likelihoods are given by:

$$\text{Prob}[\text{cc}|\text{disclose}] = p_{cc}/[p_{cc} + y_{C\text{-}Sorting} \cdot p_{FR}] \geq p_{cc} \text{ when disclosure is costly.}$$

$$\text{Prob}[\text{cc}|\text{disclose}] = p_{cc}/[p_{cc} + y_{F\text{-}Sorting} \cdot p_{FR}] \geq p_{cc} \text{ when disclosure is free.}$$

The way to derive the optimal control path of a strategic free rider i in this case is identical to the case in Section B.1.1, except that the probability of being matched with a conditional cooperator in the Hamiltonian for $t > 1$ in Condition (#0) is Prob[cc|disclose], instead of p_{cc}. By re-doing exactly the same calculation process explained in Section B.1.1, the following reputation building behaviors can be derived:

Summary 2: *If there is such a large fraction of conditional cooperators in the set of disclosers that Prob[cc|disclose] > .25, i always contributes 20 and then discloses the state in the next period until period 19 in a given sorting treatment (in period 20, i discloses the state and then* *selects $c_{i,20} = 0$). Since it is always the case that Prob[cc|disclose] $\geq p_{cc}$, when Prob[cc|disclose] > .25, the disclosure rate and average contribution amount are both higher in the C-Sorting (F-Sorting) than in the C-RM (F-RM) treatment. By contrast, if Prob[cc|disclose] < .25 in the C-Sorting treatment, i always contributes nothing and hides her state. If Prob[cc|disclose] < .25 in the F-Sorting treatment, i always contributes nothing and randomly decides whether to disclose in each period.*

B.1.3. *Mandatory treatment*

Under the assumption that no free riders other than *i* mimic the behavior of a conditional cooperator, the analysis for the Mandatory treatment is the same as the one studied in Section B.1.1. In the Mandatory treatment, there are two types as below in the population:

$p_{cc}\times100$: the percentage of those who act on the conditional cooperation strategy

$p_{FR}\times100$: the percentage of those who behave opportunistically (contribute 0)

As the likelihood of being matched with a conditional cooperator is p_{cc}, the setup with mandatory disclosure would not differ from the two random-matching treatments.

Summary 3: *Subjects' contribution behaviors in the Mandatory treatment are the same as those in the C-RM and F-RM treatments outlined in Summary 1.*

B.2. When some free riders other than *i* mimic the behavior of a conditional cooperator

If it is instead assumed that some free riders other than *i* mimic the behavior of a conditional cooperator, then *i*'s mimicking incentives change.

Assumption: $p_{m,k}\times100\%$ of free rider types mimic the behavior of a conditional cooperator in the treatment $k \in$ {C-RM, F-RM, C-Sorting, F-Sorting}. Among the rest, $(1-p_{m,k})\times100\%$ of free rider types, $y_k\times100\%$ disclose their states.

This change in the assumption alters the likelihood that free rider *i* is matched with a person who acts according to the conditional cooperation strategy (denoted *p* in Proposition B.1) as follows:

$$\begin{cases} p = p^{RM} = p_{cc} + p_{FR} \cdot p_{m,k}^{RM} \text{ for both the C-RM and F-RM treatments.} \\ p = p^{S} = \text{Prob}[cc|disclose] = \frac{p_{cc}+p_{FR}\cdot p_{m,k}^{S}}{p_{cc}+p_{FR}\cdot p_{m,k}^{S}+y_k\cdot p_{FR}\cdot(1-p_{m,k}^{S})} \text{ for the F-Sorting treatment.} \\ p = p^{M} = p_{cc} + p_{FR} \cdot p_{m,k}^{M} \text{ for the Mandatory treatment.} \end{cases}$$

Here, the superscripts, *RM*, *S* and *M*, refer to the random matching, sorting, and Mandatory conditions, respectively.

The rest is to execute exactly the same calculation process explained in Section B.1. The threshold probability for the Bang-bang solution is the same as in Section B.1 (i.e., 25%), while p^{RM}, p^{S}, and p^{M} are used as p.

References:

Dufwenberg, M. and Kirchsteiger, G. "A theory of sequential reciprocity." *Games and Economic Behavior*, Vol. 47 (2004), pp. 268-298.

Fehr, E. and Schmidt, K. "A theory of fairness, competition, and cooperation." *Quarterly Journal of Economics*, Vol. 114 (1999), pp. 817-868.

Fischbacher, U. and Gächter, S. "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments." *American Economic Review*, Vol. 100 (2010), pp. 541-56.

Fischbacher, U., Gächter, S. and Fehr, E. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics Letters*, Vol. 71 (2001), pp. 397-404.

Forsythe, R., Isaac, M. and Palfrey, T. "Theories and Tests of "Blind Bidding" in Sealed-bid Auctions." *Rand Journal of Economics*, Vol. 20 (1989), pp. 214-238.

Kamei, K. "Endogenous Reputation Formation under the Shadow of the Future." *Journal of Economic Behavior & Organization*, Vol. 142 (2017), pp. 189-204.

Rabin, M. "Incorporating Fairness Into Game Theory and Economics." *American Economic Review*, Vol. 83 (1993), pp. 1281-1302.

Sethi, S. and Thompson, G. *Optimal Control Theory* (2nd edition). Springer, 2006.

Appendix C: Additional Tables and Figures

Table C.1. *Treatment Differences in Subjects' Behaviors based on a Regression Approach*

Dependent variable: / Independent variable:	Disclosure rate [%] (1)	Avg. contribution [ECUs] (2)	Avg. Belief [ECUs] (3)	Avg. Payoff [ECUs] (4)
(#1) C-RM dummy {= 1 for the C-RM treatment; 0 otherwise}	---	1.99	1.75	.97
		(1.39)	(1.16)	(.82)
(#2) C-Sorting dummy {= 1 for the C-Sorting treatment; 0 otherwise}	.29***	5.19***	5.62***	2.60***
	(.07)	(1.38)	(1.05)	(.80)
(#3) F-RM dummy {= 1 for the F-RM treatment; 0 otherwise}	.30***	3.57**	2.91**	2.14**
	(.06)	(1.52)	(1.36)	(.91)
(#4) F-Sorting dummy {= 1 for the F-Sorting treatment; 0 otherwise}	.48***	4.07***	3.56**	2.44***
	(.05)	(1.58)	(1.41)	(.94)
(#5) Mandatory dummy {= 1 for the Mandatory treatment; 0 otherwise}	---	7.62***	7.32***	4.57***
		(1.24)	(2.47)	(1.29)
Constant	.23***	3.91***	4.12***	22.3***
	(.041)	(1.24)	(1.00)	(.74)
Number of observations	304	456	456	456
Reference group	C-RM	Baseline	Baseline	Baseline
p-value (two-sided) for Wald χ^2 tests to the following hypothesis:				
H_0: (#1) = (#2)	---	.0003***	.0000***	.0004***
H_0: (#1) = (#3)	---	.1450	.2920	.0612*
H_0: (#1) = (#4)	---	.0729*	.1187	.0290**
H_0: (#1) = (#5)	---	.0027***	.0169**	.0012***
H_0: (#2) = (#3)	.8245	.1266	.0056***	.4526
H_0: (#2) = (#4)	.0030***	.3273	.0486**	.8151
H_0: (#2) = (#5)	---	.1937	.4529	.0735*
H_0: (#3) = (#4)	.0001***	.6995	.6332	.6995
H_0: (#3) = (#5)	---	.0398**	.0702*	.0398**
H_0: (#4) = (#5)	---	.0784*	.1266	.0784*

Notes: Linear regressions. Standard errors, in parentheses, were clustered by session. Random effects were included to control for panel structure because treatment dummies are included as regressors. Session-average data were used. Observations from period 2 to 20 in the four treatments with information disclosure were used as data in column (1). Observations from period 2 to 20 in all six treatments (C-RM, C-Sorting, F-RM, F-Sorting, Baseline, Mandatory) are used in columns (2) to (4).

*, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

Table C.2. *A Comparison of Disclosers' Contribution Behaviors Across the Four Treatments with Voluntary Information Disclosure (supplementing Fig. 2 of the paper)*

The following table includes two-sided *p*-values based on Wald (Chi-squared) tests to compare disclosers' contribution behaviors or payoffs among the treatments (see Fig. 2 for the average contribution amounts and payoffs by treatment):

[For disclosers' average contributions:]

Hypothesis:	Chi-squared	Prob > Chi-squared
H_0: C-RM = C-Sorting	.39	.5301
H_0: C-RM = F-RM	.37	.5445
H_0: C-RM = F-Sorting	2.44	.1183
H_0: C-Sorting = F-RM	.85	.3558
H_0: C-Sorting = F-Sorting	3.61	.0576*
H_0: F-RM = F-Sorting	.36	.5492

There is a weakly significant difference in the disclosers' contribution amounts between the C-Sorting and F-Sorting treatments, similar to the treatment difference between the two treatments discussed in the paper. Aside from this, no other treatment differences are found.

[For disclosers' average payoffs:]

Hypothesis:	Chi-squared	Prob > Chi-squared
H_0: C-RM = C-Sorting	17.72	.0000***
H_0: C-RM = F-RM	1.97	.1600
H_0: C-RM = F-Sorting	6.12	.0134**
H_0: C-Sorting = F-RM	5.40	.0201**
H_0: C-Sorting = F-Sorting	0.34	.5604
H_0: F-RM = F-Sorting	1.50	.2204

There are significant differences in the disclosers' payoffs behaviors between the C-RM and C-Sorting treatments, between C-RM and F-Sorting treatments, and between C-Sorting and F-RM treatments.

Note: *, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

Table C.3. *Subjects' Disclosure Decisions and Last Period Experiences*

Dependent variable: A dummy variable which equals 1 if subject *i* chose to disclose her state (last-period contribution amount) in period $t \in \{2, 3, \ldots, 19\}$; and 0 otherwise.

Treatment:	(a) C-Sorting treatment			(b) C-RM treatment		
	All data	*i* disclosed in period *t* – 1	*i* did not disclose in period *t* – 1	All data	*i* disclosed in period *t* – 1	*i* did not disclose in period *t* – 1
Independent variable:	(a1)	(a2)	(a3)	(b1)	(b2)	(b3)
Period Number {= 2, 3, …, 19}	-.0021 (.0036)	-.0036 (.0055)	.0024 (.0049)	-.0076* (.0026)	-.00041 (.0062)	-.0035* (.0012)
The contribution amount of subject *i* in period *t* – 1	---	.028** (.0076)	.026 (.014)	---	.044*** (.0030)	.031*** (.0039)
Disclosed-Last-Period dummy × *i*'s period *t* – 1 matched partners' contribution amounts	---	.0023 (.0033)	---	---	.0037 (.0033)	.0010 (.0019)
Not-Disclosed-Last-Period dummy × *i*'s period *t* – 1 matched partners' contribution amounts	---	---	-.0040 (.0030)	---	.00076 (.0021)	.00010 (.0012)
Constant	.56*** (.038)	.39** (#1) (.091)	.14 (.16)	.31*** (.027)	.028 (#2) (.096)	.038 (.027)
# of observations	792	404	344	864	194	622

Notes: Linear regressions. Standard errors, in parentheses, were clustered by session. Individual fixed effects were included to control for panel structure. The Disclosed-Last-Period dummy equals 1(0) if subject *i*'s period *t* – 1 partner disclosed (did not disclose) his/her last period contribution amount in period *t* – 1. The Not-Disclosed-Last-Period dummy equals 1(0) if subject *i*'s period *t* – 1 partner did not disclose (disclosed) his/her last period contribution amount in period *t* – 1. Observations for which the period number is greater than 1 and less than 20 were included in the analyses, considering the strong end-game effect observed in the experiment. Results are similar even if data from period 20 are included (the results are omitted to conserve space). *, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

- Two-sided *t* test result for H_0: (#1) = (#2)

  $t = 2.74$, and $p = .0062$*** (i.e., Period *t* – 1 disclosers in the C-Sorting treatment were significantly more likely than disclosers in the C-RM treatment to disclose their states again in period *t*)

(Continued)

Treatment:	(c) F-Sorting treatment			(d) F-RM treatment		
	All data	*i* disclosed in period $t-1$	*i* did not disclose in period $t-1$	All data	*i* disclosed in period $t-1$	*i* did not disclose in period $t-1$
Independent variable:	(c1)	(c2)	(c3)	(d1)	(d2)	(d3)
Period Number {= 2, 3, …, 19}	.00059 (.0038)	.0016 (.0025)	.012** (.0027)	-.0017 (.0023)	.0062** (.0015)	-.0017 (.0033)
The contribution amount of subject *i* in period $t-1$	---	.022** (.0065)	.038* (.012)	---	.039*** (.0029)	.040** (.0074)
Disclosed-Last-Period dummy × *i*'s period $t-1$ matched partners' contribution amounts	---	.0049** (.0014)	---	---	-.0024 (.0014)	-.0093 (.0041)
Not-Disclosed-Last-Period dummy × *i*'s Period $t-1$ matched partners' contribution amounts	---	---	.00014 (.0019)	---	.00033 (.0023)	-.0057 (.0048)
Constant	.70*** (.039)	.55*** (#3) (.090)	.13 (.099)	.56*** (.025)	.26*** (#4) (.035)	.20*** (.030)
# of observations	792	528	220	792	411	337

- Two-sided *t* test result for H_0: (#3) = (#4)

$t = 3.00$, and $p = .0027$*** (i.e., Period $t-1$ disclosers in the F-Sorting treatment were significantly more likely than disclosers in the F-RM treatment to disclose their states again in period *t*)

Table C.4. *Subjects' Decisions to Contribute by Cooperation Type (supplementing Fig. 4 of the paper)*

Dependent variable: subject *i's* contribution amount to the joint account in period *t*.

Treatment: / Independent variable:	C-RM (1)	F-RM (2)	C-Sorting (3)	F-Sorting (4)
(a) Conditional cooperator dummy{= 1 if subject *i* is a conditional cooperator; and 0 if subject *i* is a free rider}	.91	1.42	-4.41	-.59
	(.95)	(2.50)	(4.29)	(3.02)
(b) Conditional cooperator dummy × Disclosure dummy{= 1 if subject *i* disclosed his/her last-period contribution in period *t*; and 0 otherwise}	3.32***	2.76***	2.28*	-.030
	(1.03)	(.50)	(1.22)	(1.81)
(c) Free rider dummy × Disclosure dummy	1.34	5.27***	-2.27	-2.73
	(2.79)	(1.83)	(2.97)	(2.33)
(d) Conditional cooperator dummy × Period 20 dummy{= 1 for period 20; 0 otherwise}	-1.05	-3.44***	-3.88***	-2.59*
	(1.19)	(1.06)	(1.12)	(1.35)
(e) Free rider dummy × Period 20 dummy	-4.27***	-3.78***	-13.9***	-.53***
	(.45)	(1.02)	(2.59)	(.15)
(f) Variable (b) × Period 20 dummy	.11	1.35	-.64	1.52
	(1.72)	(2.57)	(3.70)	(2.92)
(g) Variable (c) × Period 20 dummy	1.02	-2.95	5.11***	-4.78
	(2.92)	(2.02)	(1.48)	(4.23)
Constant	4.38***	4.71**	11.74***	8.98***
	(.34)	(2.36)	(3.63)	(2.47)
# of observations	646	703	570	646
• Two-sided *p*-values for Wald tests for the following comparisons regarding disclosers:				
H_0: (a) + (b) = (c) [the null that the contribution amounts of conditional cooperators are the same as those of free riders before period 20]	.4346	.7783	.9097	.3503
H_0: (a) + (b) + (d) + (f) = (c) + (e) + (g) [the null that the contribution amounts of conditional cooperators are the same as those of free riders in period 20]	.0305**	.0075***	.0216**	.0620*
• Two-sided *p*-values for Wald tests for the following comparison regarding non-disclosers:				
H_0: (a) + (d) = (e) [the null that the contribution amounts of conditional cooperators are the same as those of free riders in period 20]	.0007***	.4714	.0009***	.3583

Notes: Linear regressions. Standard errors, in parenthesis, were clustered by session. Random effects were included to control for panel structure, because dummy variables are included as independent variables. Observations of only subjects who were classified as conditional cooperators or free riders were used in the regressions. The reference group in each regression is free riders before period 20. Data from period 1 were not used as there were no disclosure decisions for subjects to make in that period (see the experimental design section of the paper).

*, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

RESULT A: Contribution decisions before period 20 (the end period)

Contribution amounts of free riders who chose to disclose are not significantly different from those of conditional cooperators who likewise chose to disclose (see the Wald test results for H_0: (a) + (b) = (c)).

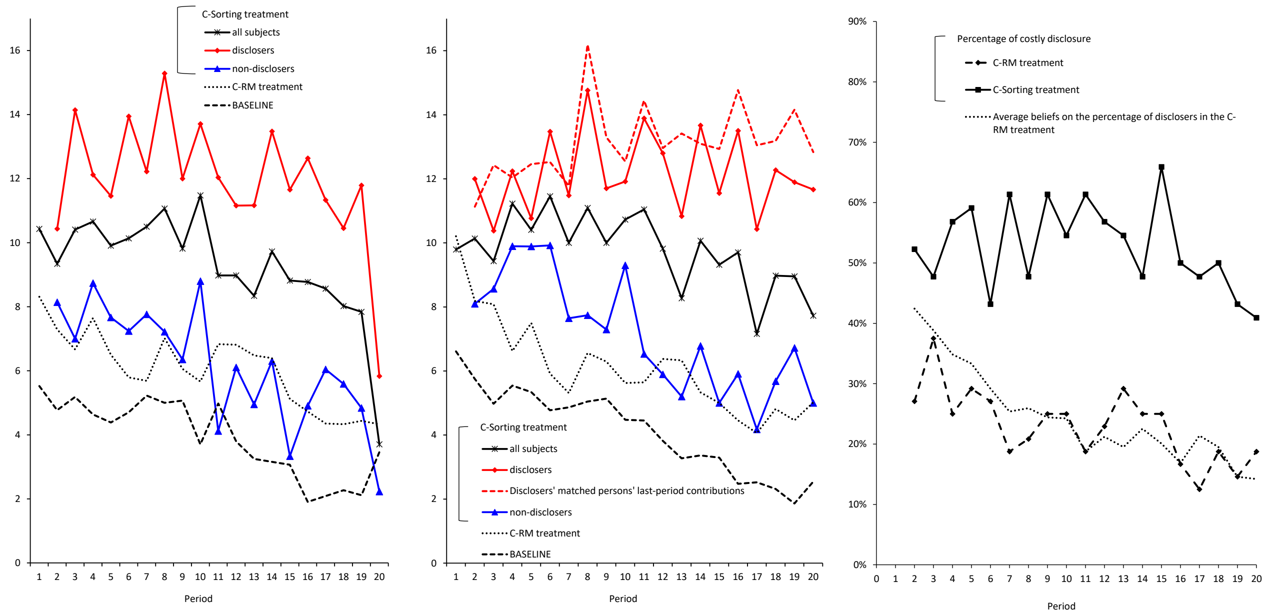
Contribution amounts of free riders who chose not to disclose are not significantly different from those of conditional cooperators who likewise chose not to disclose (see the coefficient estimates of variable (a)).

RESULT B: Contribution decisions in period 20 (the end period)

Contribution amounts of free riders who chose to disclose are significantly lower than those of conditional cooperators that likewise chose to disclose. See the Wald test results for H_0: (a) + (b) + (d) + (f) = (c) + (e) + (g).

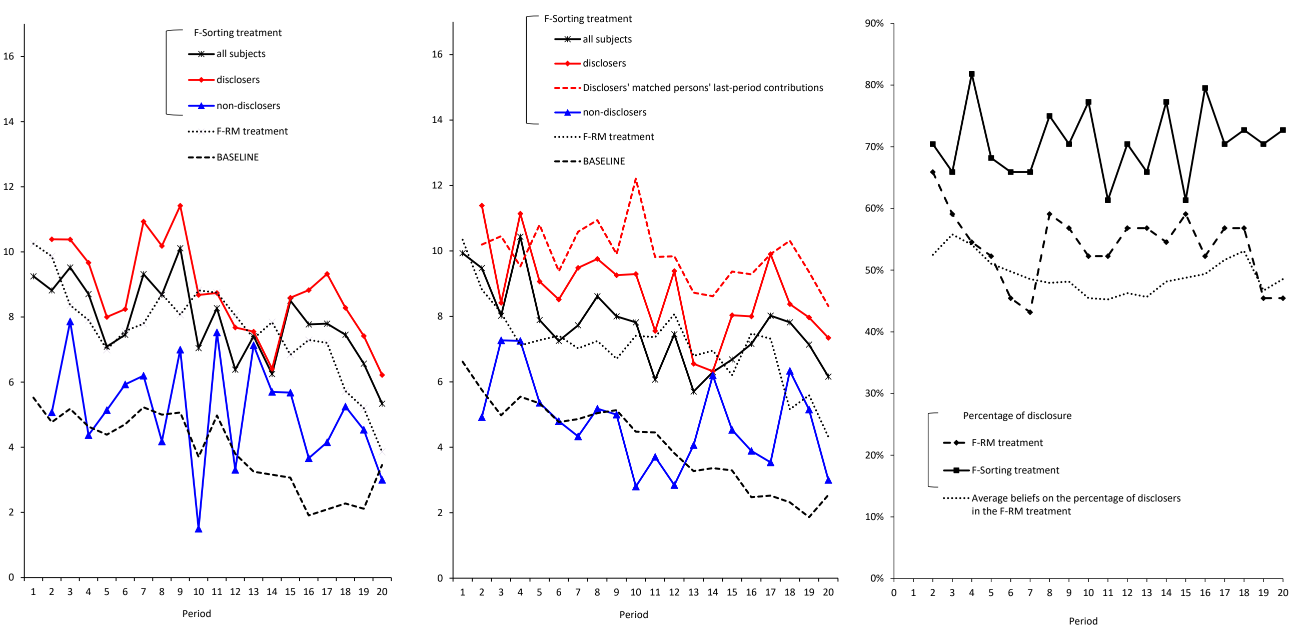
Contribution amounts of free riders who chose not to disclose are significantly lower than those of conditional cooperators that likewise chose not to disclose in treatments when disclosing was costly. See the Wald test results for H_0: (a) + (d) = (e).

Fig. C.1. *Period-by-Period Average Contributions and the Subjects' Disclosure Rates*



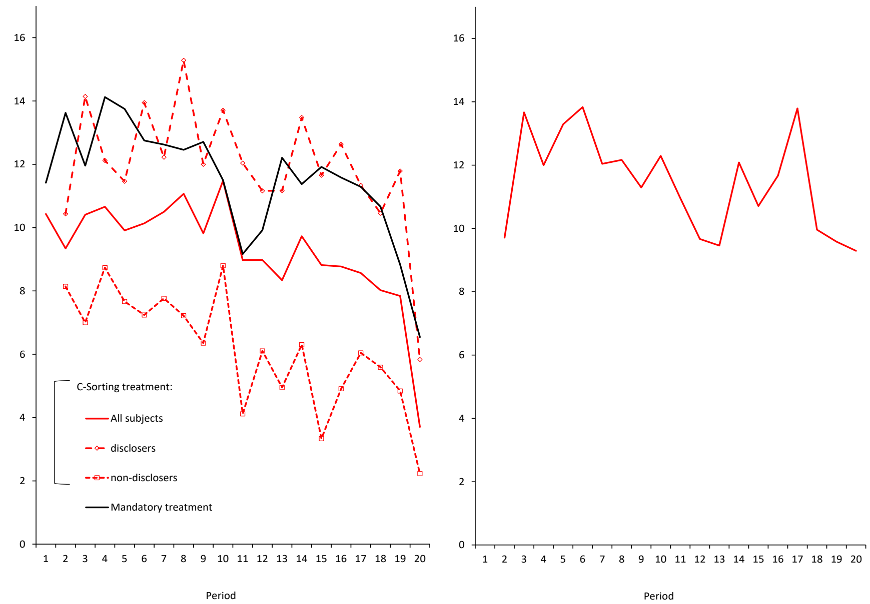
(a) Avg. contribution [ECUs] (b) Avg. belief on partners' contributions [ECUs] (c) Percentage of disclosers [%]

(I) C-RM and C-Sorting treatments



(a) Avg. contribution [ECUs] (b) Avg. belief on partners' contributions [ECUs] (c) Percentage of disclosers [%]

(II) F-RM and F-Sorting treatments



(a) Avg. contribution [ECUs] (b) Avg. belief on partners' contributions [ECUs]

(III) C-Sorting and Mandatory treatments

Note: The red dashed lines depicted in panels I(b) and II(b) show the average of last-period contributions made by the disclosers' matched partners to the joint accounts in the C-Sorting and F-Sorting treatments, respectively.

Fig. C.2. *Cumulative Distribution of Individual Contribution Decisions by Treatment*

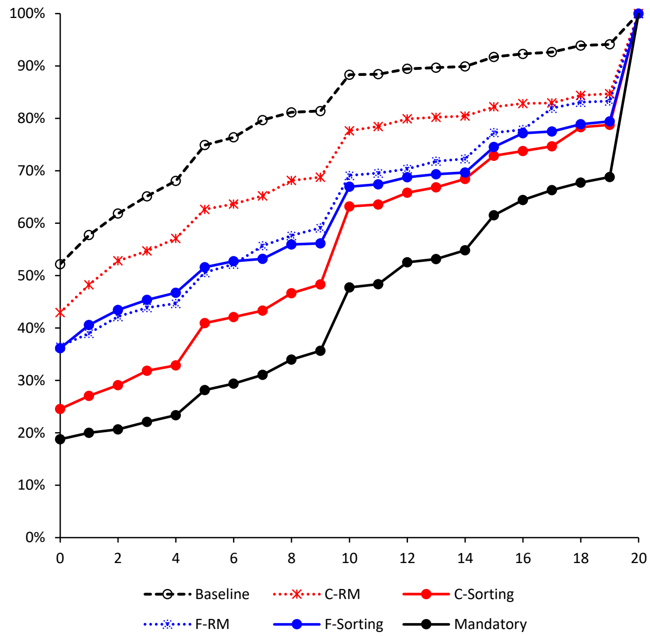
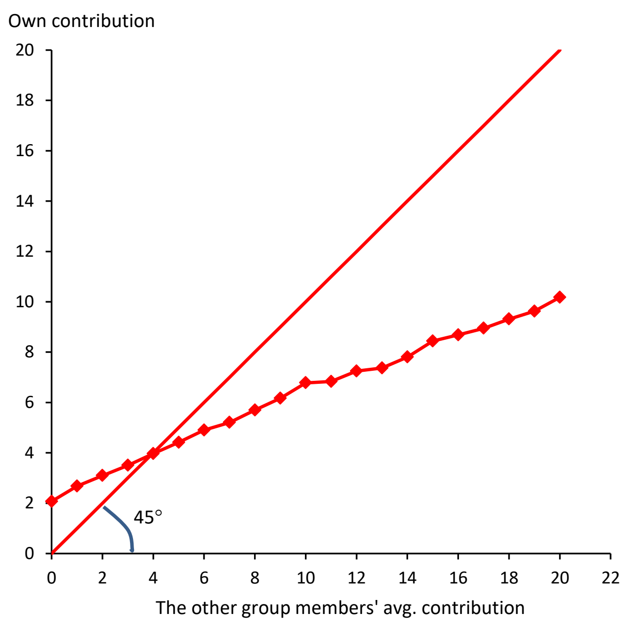


Fig. C.3. *Average Conditional Contribution Schedule*

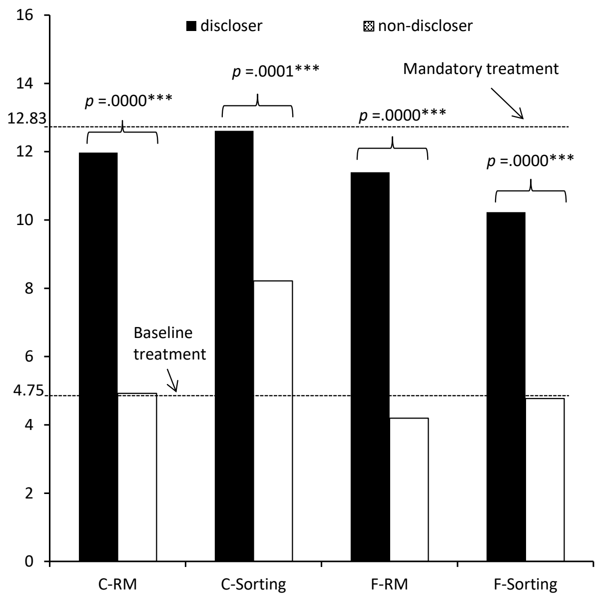


Notes: The above schedule is the average of all subjects' conditional contribution schedules.

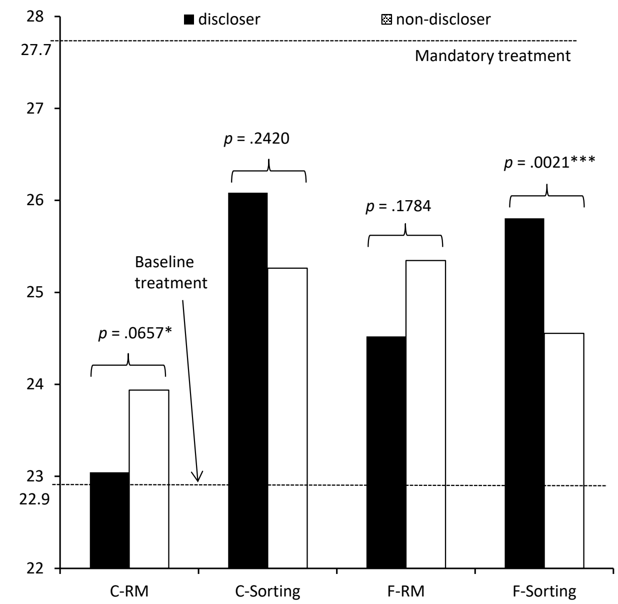
Remark: (1) The spearman's ρ correlation coefficient between own contribution amounts (y-axis) and the other group members' average contribution amounts (x-axis) is 1.0000 (two-sided p-value $< .001$).

(2) The slope, based on a linear regression, of the average conditional contribution schedule is .39, which is significantly positive (two-sided p-value $< .001$) and is significantly less than 1 (F-test, two-sided p-value $< .001$). The intercept is 2.41 ECUs, which is significantly positive (F-test, two-sided p-value $< .001$).

Fig. C.4. *Subjects' Average Contributions and Payoffs by Disclosure Decision in the First and Second Halves of the Experiment (supplementing Fig. 2 of the paper)*

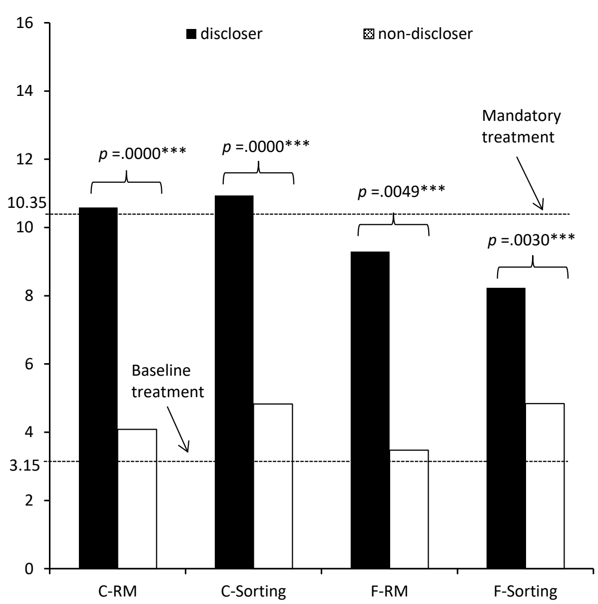


(a) Average contribution [ECUs]

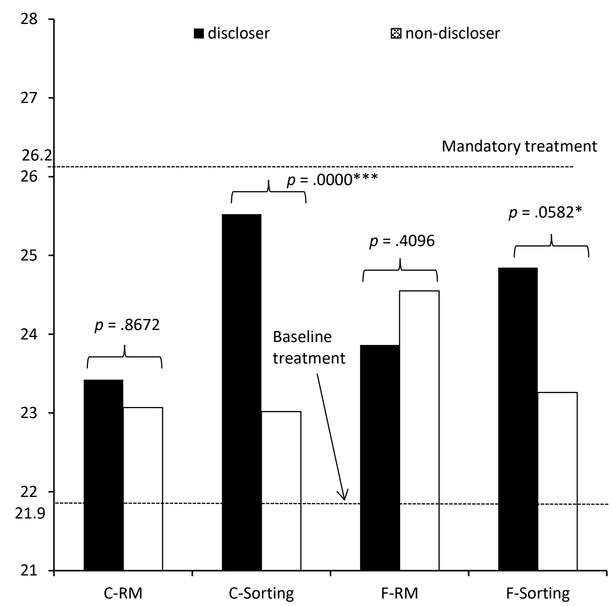


(b) Average payoff [ECUs]

(1) The first half (periods 2 to 10)



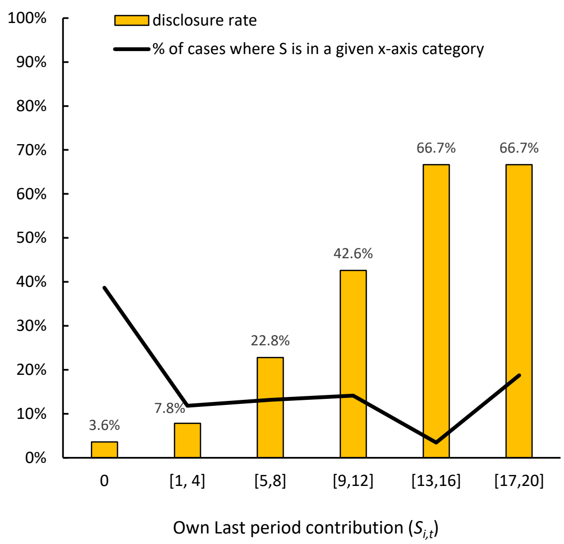
(a) Average contribution [ECUs]



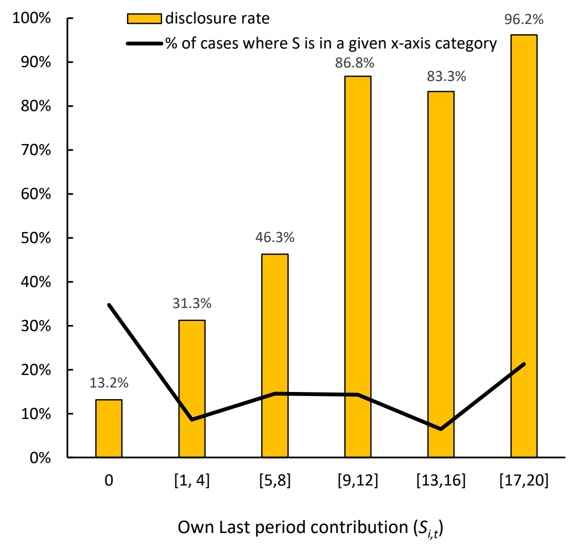
(b) Average payoff [ECUs]

(2) The second half (periods 11 to 20)

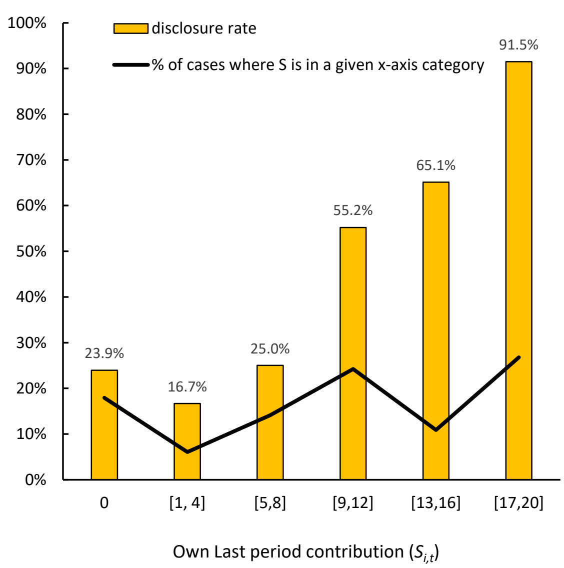
Fig. C.5. *Subjects' Disclosure Rates by the Size of their State*



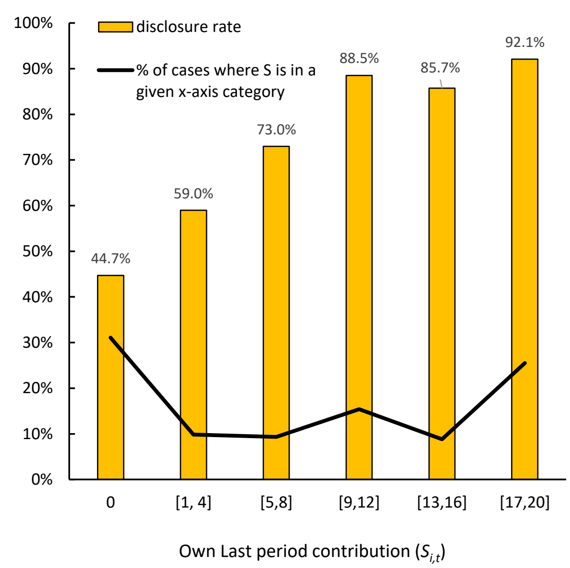
(a) The C-RM treatment



(b) The F-RM treatment

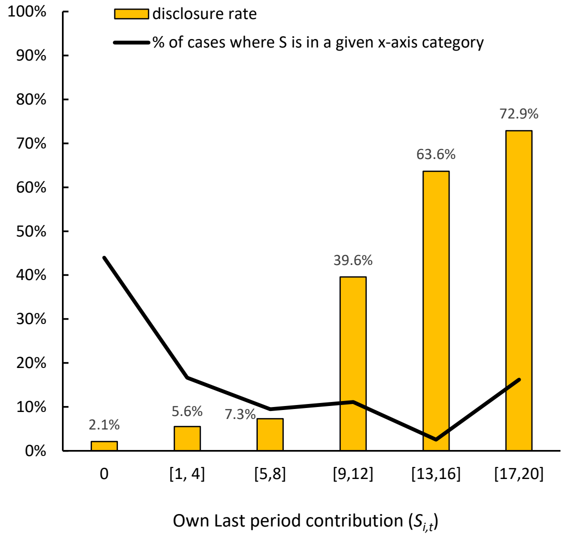


(c) The C-Sorting treatment

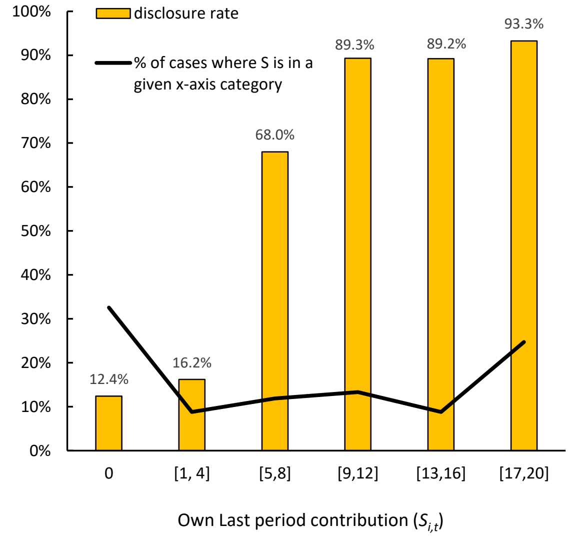


(d) The F-Sorting treatment

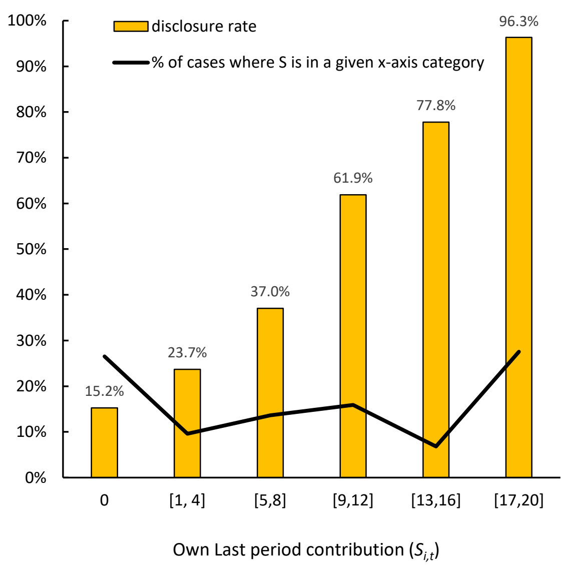
(1) The first half (periods 2 to 10)



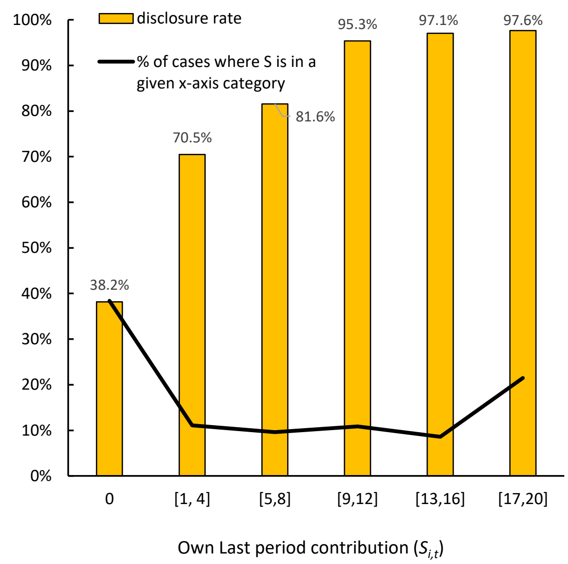
(a) The C-RM treatment



(b) The F-RM treatment



(c) The C-Sorting treatment

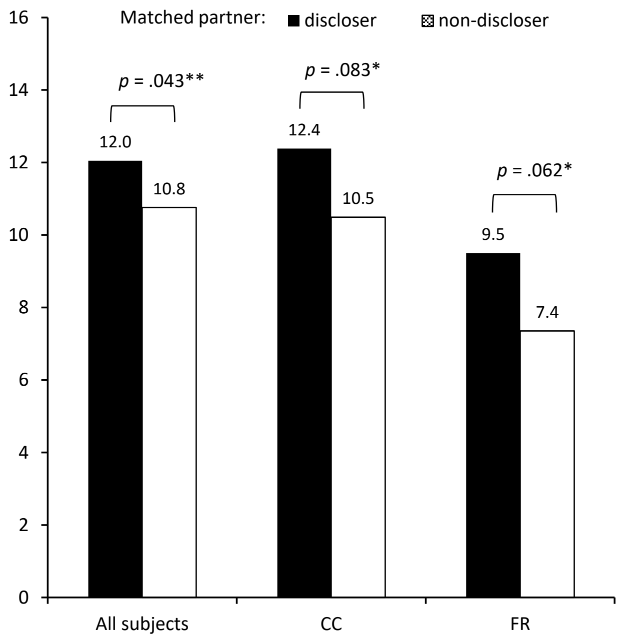


(d) The F-Sorting treatment

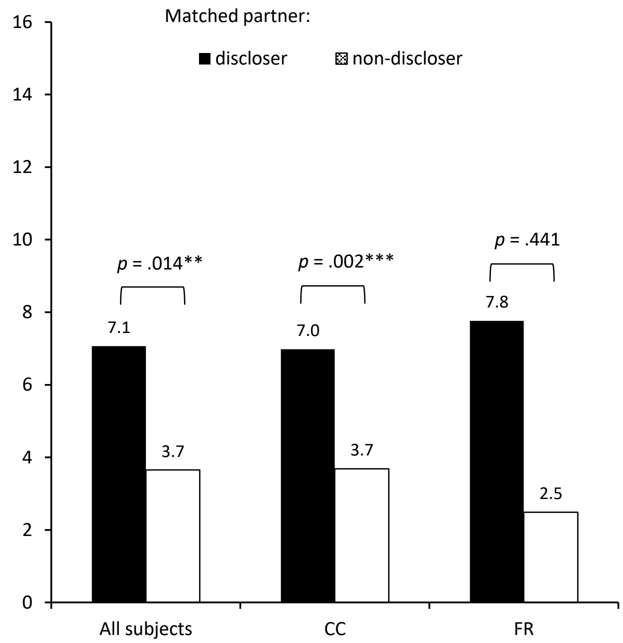
(2) The second half (periods 11 to 19)[#1]

Note: [#1] As in Fig. 3 of the paper, data from period 20 were not used. However, results are similar even if the disclosure decisions in period 20 are included (the results are omitted to conserve space).

Fig. C.6. *Average Contribution Amounts by Disclosure Decision and by Conditional Contribution Type (supplementing Fig. 4 of the paper)*

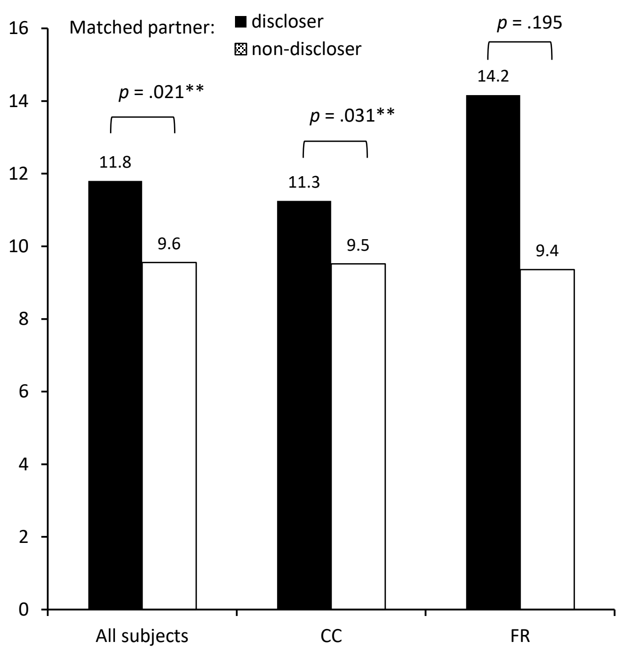


(a) Disclosers' average contribution [ECUs]

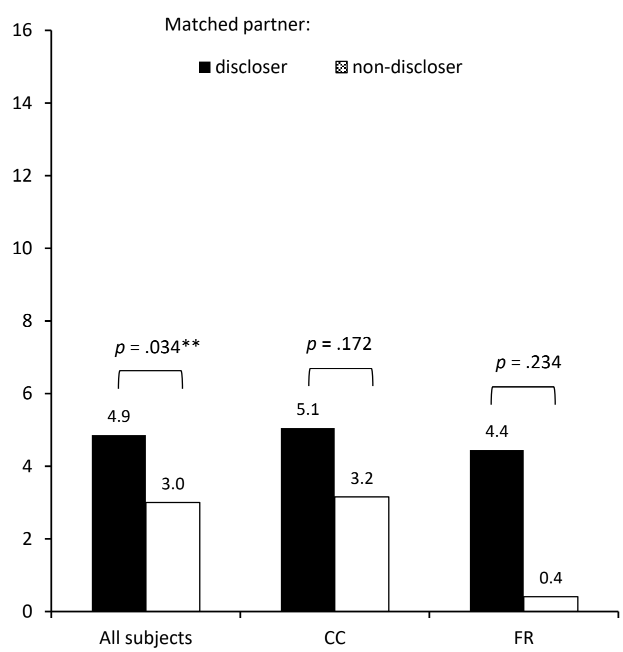


(b) Non-disclosers' average contribution [ECUs]

(1) The C-RM treatment



(a) Disclosers' average contribution [ECUs]

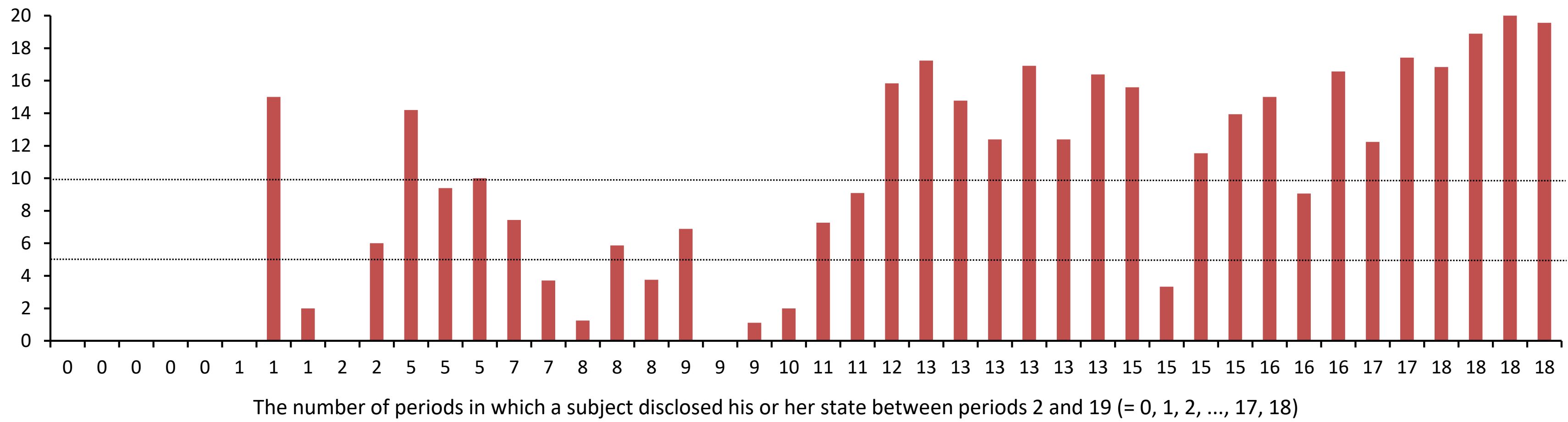


(b) Non-disclosers' average contribution [ECUs]

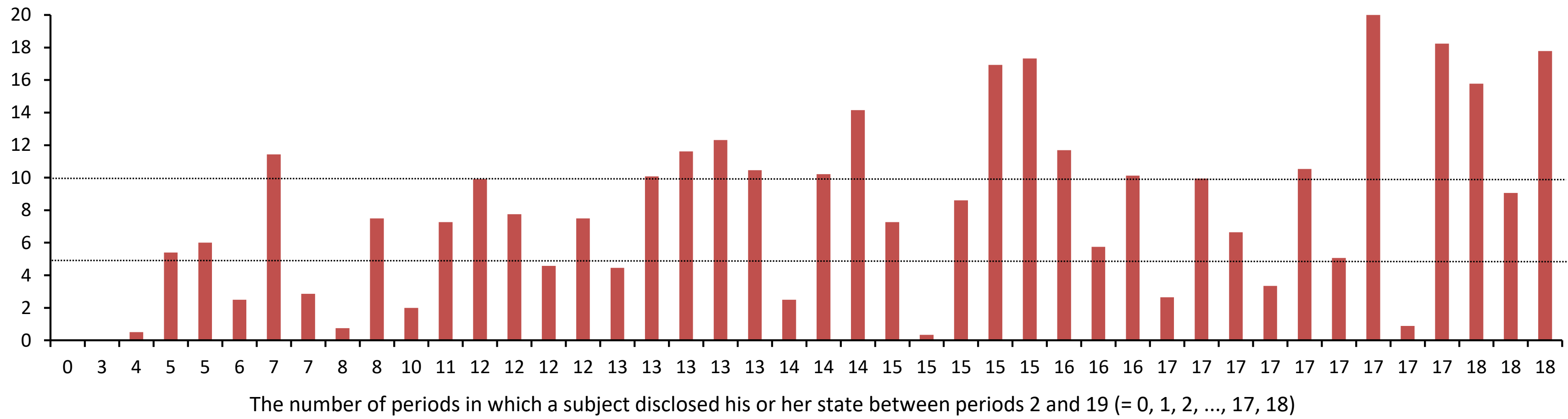
(2) The F-RM treatment

Notes: Average across all periods before period 20. *p*-values (two-sided) in the figure are based on individual fixed effect linear regressions with robust standard errors clustered by session.

Fig. C.7. *Average Contribution Amounts of Disclosers by the Total Number of Periods in which they Disclosed the States*



(1) The C-Sorting treatment



(2) The F-Sorting treatment

Note: Each bar indicates each subject's average contribution when he or she disclosed the state.

Fig. C.8. *Subject-by-Subject Patterns of Disclosure Decisions in the C-Sorting and F-Sorting treatments (supplementing Fig. C.7 of the Appendix)*

period	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
s1	0	1	0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	0
s2	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
s3	1	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	0
s4	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0
s5	1	1	1	0	0	1	0	1	0	1	1	0	1	1	1	1	1	1	1
s6	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
s7	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	1	1	0
s8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
s9	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
s10	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
s11	1	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0	1	0	1
s12	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	0	1	1
s13	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0
s14	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s15	0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	1	1
s16	0	0	1	0	1	0	1	1	0	1	1	0	1	1	0	1	0	1	1
s17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s18	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
s19	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
s20	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0
s21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1
s22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
s23	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	1
s24	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	1
s25	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
s26	0	1	0	1	1	0	1	0	1	0	0	0	1	1	1	1	1	1	1
s27	0	1	1	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1
s28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
s29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
s30	0	1	1	0	1	1	0	1	0	1	0	1	1	0	1	0	1	1	1
s31	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
s32	1	0	1	1	1	1	1	0	1	1	1	1	1	0	0	1	0	1	0
s33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s34	1	1	0	0	1	0	1	1	1	0	1	1	0	0	1	1	0	0	0
s35	1	1	0	0	1	0	1	0	1	1	0	1	1	1	0	1	0	0	1
s36	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1
s37	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0
s38	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s40	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
s41	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
s42	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
s43	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
s44	0	1	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1	1

(I) C-Sorting treatment

period	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
s1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1
s2	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1
s3	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
s4	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1
s5	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	1	1	0	0
s6	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0
s7	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
s8	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
s9	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1
s10	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s11	1	1	0	1	0	1	1	0	0	1	1	1	0	1	0	1	0	0	0
s12	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
s13	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
s14	1	1	0	0	0	0	0	1	0	0	0	1	0	1	0	0	1	0	1
s15	0	0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
s16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s17	1	1	1	1	1	1	0	0	1	1	1	0	0	1	1	0	1	1	0
s18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
s19	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
s20	0	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	1
s21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s22	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
s23	0	1	1	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0
s24	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
s25	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0
s26	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0
s27	1	1	0	0	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0
s28	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
s29	0	1	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0
s30	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
s31	1	1	0	1	1	0	1	0	1	1	0	1	1	0	0	1	0	0	0
s32	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s33	1	1	1	0	0	0	1	1	0	1	1	1	0	1	1	0	0	1	0
s34	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	1	0	1
s35	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
s36	0	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1
s37	0	1	0	1	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0
s38	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
s39	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
s40	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	1	0	1
s41	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
s42	0	0	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1
s43	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
s44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(II) F-Sorting treatment

Notes: 1 (0) refers to a case in which a subject did not disclose (disclosed) the state. Cases of non-disclosure are highlighted in yellow.

Remark: These two panels reveal different disclosure patterns between the C-Sorting and F-Sorting treatments. Panel I suggests that although some subjects switched back and forth between disclosing and not disclosing, those who constantly disclosed as well as those who constantly hid their behaviors are prevalent in the C-Sorting treatment. By contrast, Panel II shows that in the F-Sorting treatment, the history of hiding and also disclosing spread across almost all subjects. Most subjects switched between disclosing and not disclosing with some duration.

This observation can be partially confirmed by a regression analysis in which the dependent variable is subject *i*'s decision to disclose (=1 if *i* discloses; 0 otherwise) and the independent variables are (a) a dummy which equals 1(0) if *i* disclosed (did not disclose) in period $t-1$, (b) the C-Sorting dummy which equals 1(0) for the C-Sorting (F-Sorting) treatment, and (c) the interaction term between variable (a) and variable (b). Observations in the C-Sorting and F-Sorting treatments were used for this analysis. According to a random effect linear regression with robust standard errors clustered by session, variable (c) has a significantly positive coefficient estimate at the 10% level. This suggests that subjects who disclosed in period $t-1$ in the C-Sorting treatment are more likely to continue to be a discloser in period *t*, compared with those in the F-Sorting treatment.

Appendix D: Analysis using Elicited Beliefs

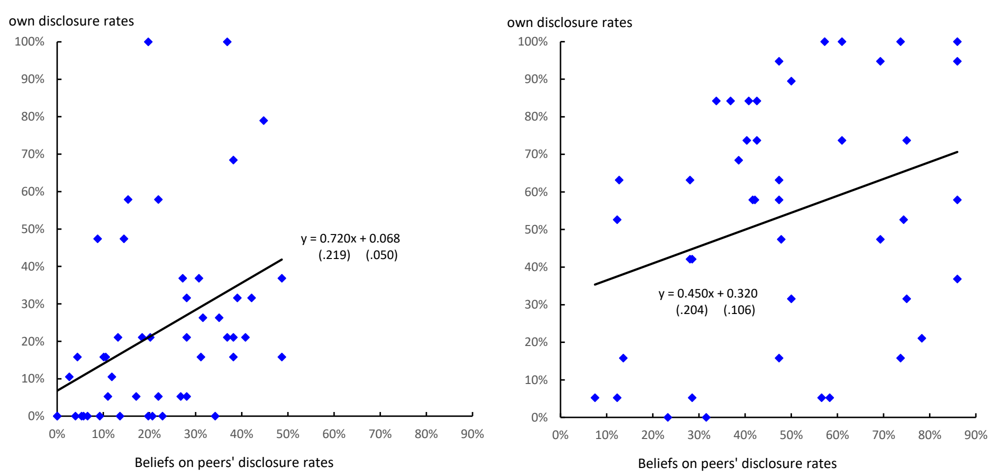
D.1. Conditional Disclosure Behaviors under Random Matching

Analyses in Section 4 of the paper largely confirmed the implication based on conditional cooperation (Appendix B) that a strategic free rider would have material incentives to mimic the behavior of conditional cooperator, provided that *p* defined in Appendix B.1 is sufficiently high (e.g., Result 1(ii), Result 2(i), and Result 3(ii) of the paper). The importance of *p* can be checked further using the data of the two random-matching treatments. Disclosers in the C-RM and F-RM treatments were not assured to be matched with disclosers. If the above discussion is appropriate, subjects' likelihoods of disclosing their states $S_{i,t}$ (last-period contribution amounts) in random-matching treatments should depend on their beliefs on *p*.

Consistent with this conjecture, subjects' average disclosure rates were significantly positively correlated with their beliefs regarding the peers' disclosure rates (Fig. D.1). Thus, the model of subjects' conditional cooperation strategies can help explain the subjects' average disclosure behaviors. The subjects' conditional disclosure behaviors also resonate with the idea that subjects would positively respond to peers' signals of future cooperation. People's conditional behaviors have been observed in various contexts, including direct punishment of defectors (Kamei, 2014), third party punishment (Kamei, forthcoming), and costly gossiping (Kamei and Putterman, 2018).

Result D.1: *The larger the fraction of peers that subjects believed would disclose in the C-RM and F-RM treatments, the more likely they themselves were to disclose the states in those treatments.*

Fig. D.1. *Subjects' Average Disclosure Rates and Beliefs on the Peers' Disclosure Behaviors*



(a) The C-RM treatment (b) The F-RM treatment

Notes: Each point indicates each subject's observation. The numbers in parenthesis in the linear equations (OLS) are robust standard errors. The slopes in panels (a) and (b) are significant at the 1% and 5% levels, respectively.

D.2. Subjects' Responses to Peers' Disclosure Decisions

Some subjects' hiding behaviors seen in Fig. 3 and Result 4 of the paper may be explained by their partners' responses to non-disclosure or peers' responses to disclosers' reputation states. For example, a discloser may contribute a large amount even toward a non-discloser to maintain her high reputation state. This could be a force that makes hiding behaviors materially beneficial under random matching as seen in Result 3(ii) of the paper. Yet understanding the subjects' behaviors is not that simple.

First, subjects' contribution amounts were positively correlated with their beliefs for both disclosers and non-disclosers in all treatments (Table D.1(a)).[7,8] Second, independent of the treatment condition, the subjects' beliefs were positively correlated with the partners' states $S_{j,t}$ when those states were observable (Table D.1(b)). The subjects' reluctance to disclose when $S_{i,t}$ was low in the C-RM treatment can be partially explained by the positive disclosure cost. In this treatment, the subjects believed that masked partners would make contribution decisions indistinguishably from disclosers whose states $S_{j,t}$ were zero (see the Partner-disclose dummy in Table D.1(b1)). However, if partner j disclosed and $S_{j,t} > 0$, subject i will believe that j would contribute more than a non-discloser, and this belief is increasing in the size of $S_{j,t}$. The estimates in Table D.1 suggest that the material benefit from disclosure exceeds the cost of one ECU for $S_{j,t} > 7.27$.[9] In fact, subjects' disclosure rates were low for all $S \leq 8$ (see Fig. 3(a) of the paper).

By contrast, hiding behaviors for subjects with $S_{i,t} > 0$ in the F-RM treatment cannot be explained by material considerations. As in the C-RM treatment, these subjects treated masked partners as if they were disclosers with $S_{j,t} = 0$ (see again Table D.1(b1)). Thus, the masked subjects who had positive $S_{j,t}$ could have obtained higher payoffs by disclosing the states and contributing zeros to maximize payoffs. Some subjects might not have recognized the benefit of disclosure, because their peers contributed non-negligible amounts on average even towards non-disclosers to keep high reputation states.

Subjects' decisions not to disclose in the C-Sorting and F-Sorting treatments are also difficult to rationalize with pecuniary considerations alone. A subject i's partner would on average contribute $9.48+.216\times S_{i,t}$ ($6.51+.231\times S_{i,t}$) when i discloses $S_{i,t}$ in the C-Sorting (F-Sorting) treatment.[10] Because non-disclosers' average contribution was 6.15 in the C-Sorting treatment, it would be materially beneficial for i to disclose the state, even if $S_{i,t} = 0$ (notice that (9.48–

[7] Additional regressions were conducted by also including a dummy which equals 1 if i's period t partner did not disclose his state as an independent variable. The result shows that the dummy fails to obtain a significant coefficient, whereas variable (i) in Table D.1 has a significantly positive coefficient in every column. The results are omitted to conserve space.

[8] Observations in period 20 were not included in the regressions since the standard end-game effect was observed. Results are similar even if observations in period 20 are included (the results are omitted to conserve space).

[9] If the parameters in the table (.43 in panel (a) and .32 in panel (b1)) are used as lower estimates of $\partial C_{i,t}/\partial C_{i,j,t}^{b}$ and $\partial C_{i,j,t}^{b}/\partial S_{j,t}$, respectively, an expectation of i's additional contribution is calculated as $.32\times.43\times S_{j,t}$ if $S_{j,t}$ is observable.

[10] This can be calculated by $6.77+.45\times(6.03+.48\times S_{i,t})$ in the C-Sorting and $4.03+.55\times(4.51+.42\times S_{i,t})$ in the F-Sorting treatment using the coefficient estimates in panels (a) and (b2) of Table D.1.

6.15)×0.8 > 1). Likewise, disclosure was a materially beneficial action in the F-Sorting treatment because non-disclosers' average contribution was 5.50, and 6.51 +.231×$S_{i,t}$ > 5.50 always.

Table D.1. *Subjects' Contributions, Beliefs, and Partners' States*

(a) Relationship between the subjects' contributions and beliefs on partners' contributions

Dependent variable: Subject i's contribution to her joint account in period t, where $2 \leq t \leq 19$.

Matching:	Random matching				Sorting			
Decision-maker i:	Discloser		Non-discloser		Discloser[#1]		Non-discloser[#2]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Independent variable:	C-RM	F-RM	C-RM	F-RM	C-Sorting	F-Sorting	C-Sorting	F-Sorting
(i) i's belief on partner j's	.43**	.50***	.65***	.44**	.45***	.55***	.39***	.26
contribution in period t ($C^b_{i,j,t}$)	(.083)	(.076)	(.061)	(.11)	(.046)	(.078)	(.045)	(.18)
Constant	7.65***	6.42***	1.02**	1.69*	6.77***	4.03***	3.58***	4.36**
	(.66)	(.65)	(.32)	(.59)	(.56)	(.69)	(.32)	(.84)
# of observations	201	431	663	361	423	559	369	233
F	26.48	43.70	114.20	16.78	93.57	50.69	77.60	2.25
Prob > F	.0142	.0070	.0018	.0263	.0023	.0057	.0031	.2302

(b) Relationship between the subjects' belief formation and the partners' states

(b1) The C-RM and F-RM treatments

Dependent variable: Subject i's belief on her partner j's contribution in period t, where $2 \leq t \leq 19$ ($C^b_{i,j,t}$).[#3]

Decision-maker i:	Discloser		Non-discloser	
Independent variable:	(1) C-RM	(2) F-RM	(3) C-RM	(4) F-RM
Partner-disclose dummy {=1 if i's period t	1.66	-.56	-.55	-.22
partner j disclosed his state in period t}	(2.38)	(.78)	(.72)	(.56)
Partner-disclose dummy × j's state [i.e.,	.33**	.46**	.32***	.34**
period t – 1 contribution amount]	(.094)	(.10)	(.041)	(.098)
Constant	6.50***	5.76***	4.33***	3.10***
	(.26)	(.38)	(.041)	(.42)
# of observations	201	431	663	361
F	89.93	32.51	845.06	47.05
Prob > F	.0021	.0093	.0001	.0054

(b2) The C-Sorting and F-Sorting treatments

Dependent variable: Discloser i's belief on her matched discloser j's contribution in period t, where $2 \leq t \leq 19$.[#4]

	C-Sorting	F-Sorting
Independent variable:	(1)	(2)
Subject i's period t matched discloser j's state [i.e., period	.48***	.42***
t – 1 contribution amount]	(.053)	(.031)
Constant	6.03***	4.51***
	(.74)	(.33)
# of observations	384	520
F	82.56	181.71
Prob > F	.0028	.0009

Notes: Subject fixed effect linear regressions. Standard errors in parenthesis were clustered by session. Results are similar when subject random effect ordered probit regressions are instead used (the results are omitted to conserve space). [#1] A small number of the disclosers were matched with non-disclosers. Results in columns (5) and (6) are similar even if these observations are excluded. [#2] A small number of the non-disclosers were matched with disclosers. Results in columns (7) and (8) are similar even if these observations are excluded. [#3] The reference group in panel (b1) is those whose matched partners did not disclose the states. [#4] Only observations in which discloser i was matched with another discloser were used in panel (b2). *, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

Result D.2: *(i) Subjects' contribution amounts were significantly positively correlated with their beliefs regarding their partner's contribution amounts. (ii) The subjects' beliefs were significantly positively correlated with the partners' states when these states were observable. (iii) In the C-RM and F-RM treatments, subjects believed that non-disclosers' contribution behaviors would be the same as those of disclosers whose states $S_{i,t}$ were zero. (iv) In the C-Sorting and F-Sorting treatments, some subjects did not disclose even though disclosing the states was a materially beneficial action.*

D.3. Other Analyses using Elicited Beliefs

Table D.2. *Subjects' Disclosure Decisions and Beliefs on the Matched Partners' Contribution Amounts in the C-Sorting and F-Sorting treatments.*

Dependent variable: Subject *i*'s session average belief on her partners' contribution amounts to the joint accounts when *i* disclosed her state (in columns (1) and (3)) or did not disclose her state (in columns (2) and (4)) up to period 19

	C-Sorting treatment		F-Sorting treatment	
	i disclosed	*i* did not disclose	*i* disclosed	*i* did not disclose
Independent variable:	(1)	(2)	(3)	(4)
The total number of periods in which subject *i* disclosed up to period 19	.37** (.15)	.084 (.14)	.17 (.16)	.041 (.19)
Constant	7.17*** (1.76)	6.81*** (1.48)	6.37*** (2.24)	4.44* (2.50)
# of observations	39[#1]	40[#2]	43[#3]	41[#4]
F	6.53	.35	1.09	.05
Prob > F	.0148**	.5551	.3028	.8306
R-squared	.1271	-.0168	.0021	-.0244

Notes: Linear regressions. Observations in period 20 were not included as in other regression analyses (except the one with treatment effects), because of the strong end-game effect seen in the experiment.

[#1] Five subjects had never disclosed their states until period 19 in the C-Sorting treatment; thus, the total number of observations is 39 (= 44 – 5).

[#2] Four subjects had always disclosed their states until period 19 in the C-Sorting treatment; thus, the total number of observations is 40 (= 44 – 4).

[#3] One subject had never disclosed his/her state until period 19 in the F-Sorting treatment; thus, the total number of observations is 43 (= 44 – 1).

[#4] Three subjects had always disclosed their states until period 19 in the F-Sorting treatment; thus, the total number of observations is 41 (= 44 – 3).

*, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

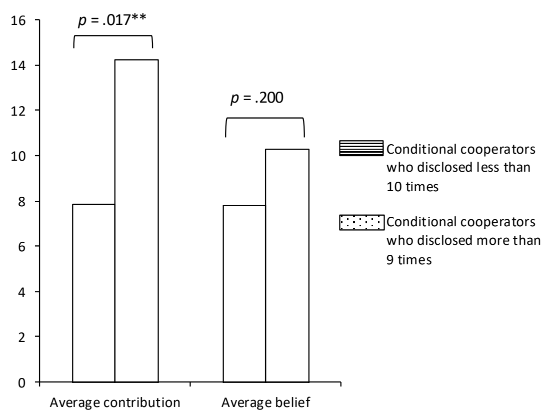
The result in the C-Sorting treatment (those who less frequently disclosed formed comparatively lower beliefs on their matched peers' contribution behaviors, as seen in column (1)) is suggestive only, because the "total number of periods in which subject *i* disclosed up to period 19" variable does not obtain a significant coefficient once session clustering is added into the regression.

Fig. D.2. *Conditional Cooperators' Average Contributions and Beliefs when They Chose to Disclose, by the Frequency of Disclosure*

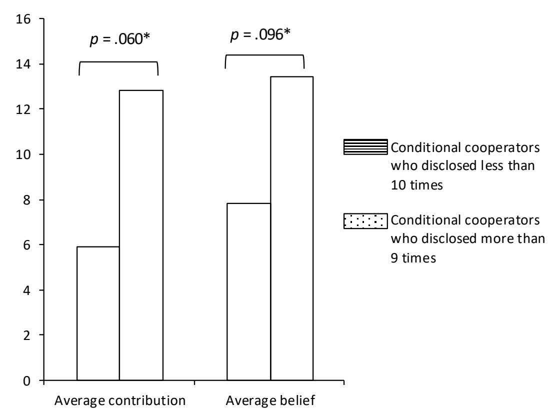
Conditional cooperators were first classified into (a) those who chose to disclose more than nine times and (b) those who chose to disclose less than ten times in the experiment. After that, the average contributions and beliefs were calculated for each category.

As shown below, conditional cooperators who less frequently disclosed formed lower beliefs on their matched peers' contribution behaviors, compared with the same types who frequently disclosed, although the differences in beliefs are not significant except for in the C-Sorting treatment.

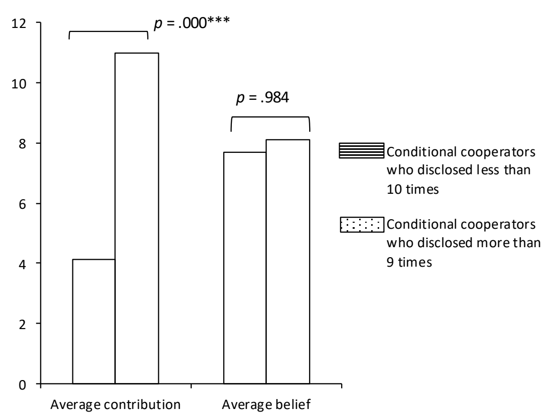
Among conditional cooperators, infrequent disclosers contributed significantly less than frequent disclosers, when they selected to disclose. This suggests that some conditional cooperators attempted to exploit other cooperators, instead of encouraging others to select conditional cooperation strategies. It follows that there are perhaps other aspects in human cooperativeness that are not captured by the classification method of Fischbacher *et al.* (2001).



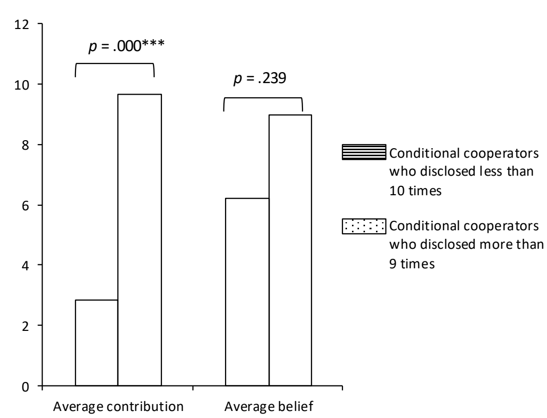
(a) C-RM treatment



(b) C-Sorting treatment



(c) F-RM treatment



(d) F-Sorting treatment

Notes: The figures show the average contributions and beliefs of conditional cooperators when they disclosed their states. The *p*-values in the figure are two-sided *p*-values, based on linear regressions (with robust standard errors clustered by session), where independent variables include a dummy which equals 1(0) if a subject disclosed (did not disclose) more than nine times in total in the experiment. In the regression, random effects were included to control for panel structure as in other regressions with dummy variables in this paper and in the Appendix.

References:

Fischbacher, U., Gächter, S. and Fehr, E. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics Letters*, Vol. 71 (2001), pp. 397-404.

Kamei, K. "Conditional Punishment." *Economics Letters*, Vol. 124 (2014), pp. 199-202.

Kamei, K. "Group Size Effect and Over-Punishment in the Case of Third Party Enforcement of Social Norms." *Journal of Economic Behavior & Organization* (forthcoming, https://doi.org/10.1016/j.jebo.2018.04.002).

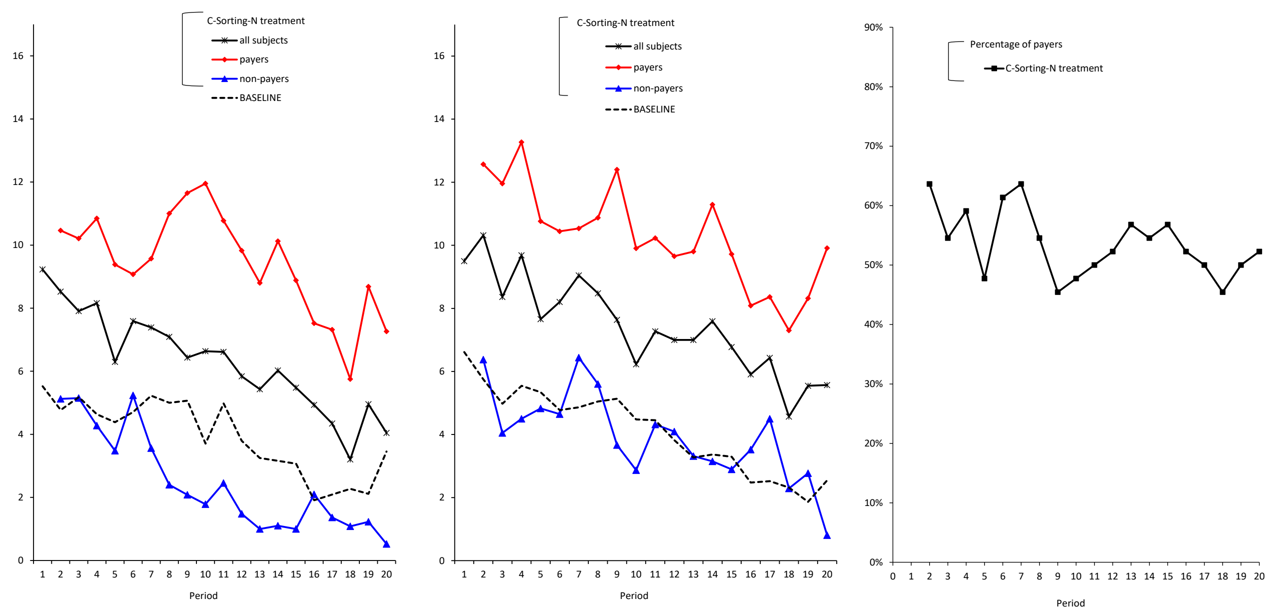
Kamei, K. and Putterman, P. "Reputation Transmission without Benefit to the Reporter: a Behavioral Underpinning of Markets in Experimental Focus." *Economic Inquiry*, Vol. 56 (2018), pp. 158-172.

Appendix E: Results of the C-Sorting-N treatment

The C-Sorting treatment achieved the highest efficiency, driven by the disclosers' strong contribution behaviors, among the four main treatments (e.g., Fig. 1 of the paper). This part of the Appendix is devoted to a further analysis for this observation.

The high performance of the C-Sorting treatment can be due to two effects: (A) the impact of disclosed information and (B) the impact of the *cost* to be matched with like-minded others. Effect (B) is absent in the F-Sorting treatment. If the presence of a positive cost is the key factor for the superior performance of the C-Sorting treatment, the disclosers' benefits from sorting may remain high even if they do not have disclosed information. As discussed in Section 2 of the paper, the "C-Sorting-N" treatment was conducted to study the relative importance of these two effects. In this treatment, subjects who paid one ECU were matched with another who paid it. However, their last-period contributions were *not* revealed. On average 53.6% of subjects paid the fee. As in the C-Sorting treatment, the average contributions were clearly different between the payers and non-payers (Fig. E.1). The difference is significant ($p < .0001$, two-sided Wald test) and large, around 7 ECUs. On average, the contributions of the payers were not significantly different from those of the disclosers in the C-Sorting treatment (Fig. E.2).

Fig. E.1. *Contributions and the Percentage of the Payers in the C-Sorting-N treatment*



(a) Avg. contribution [ECUs] (b) Avg. belief on partners' contributions [ECUs] (c) Percentage of disclosers [%]

Nevertheless, the average contribution is 2.51 ECUs lower for the payers than for the disclosers in the C-Sorting treatment (Fig. E.2). In addition, despite the lack of significant difference in the average contribution, the average payoff is significantly lower for the payers than for the disclosers in the C-Sorting treatment (see again Fig E.2). This suggests that disclosed information is essential for a payer to induce her matched payer to contribute large amounts.

As for the treatment differences, the average contribution in the C-Sorting-N treatment is 6.31 ECUs, which is around 61% higher than that in the Baseline treatment. However, the former is significantly lower than that in the C-Sorting treatment (column (1) in Table E.1). Qualitatively the same difference is found when average payoffs are compared between the C-Sorting-N and C-Sorting treatments (column (2) in Table E.1).

In short, there was an effect of having a positive disclosure cost; however, even accounting for this, the presence of disclosed information remained crucial for the high performance in the C-Sorting treatment.

Result E.1: *The average contribution and average payoff were both significantly lower in the C-Sorting-N than in the C-Sorting treatment.*

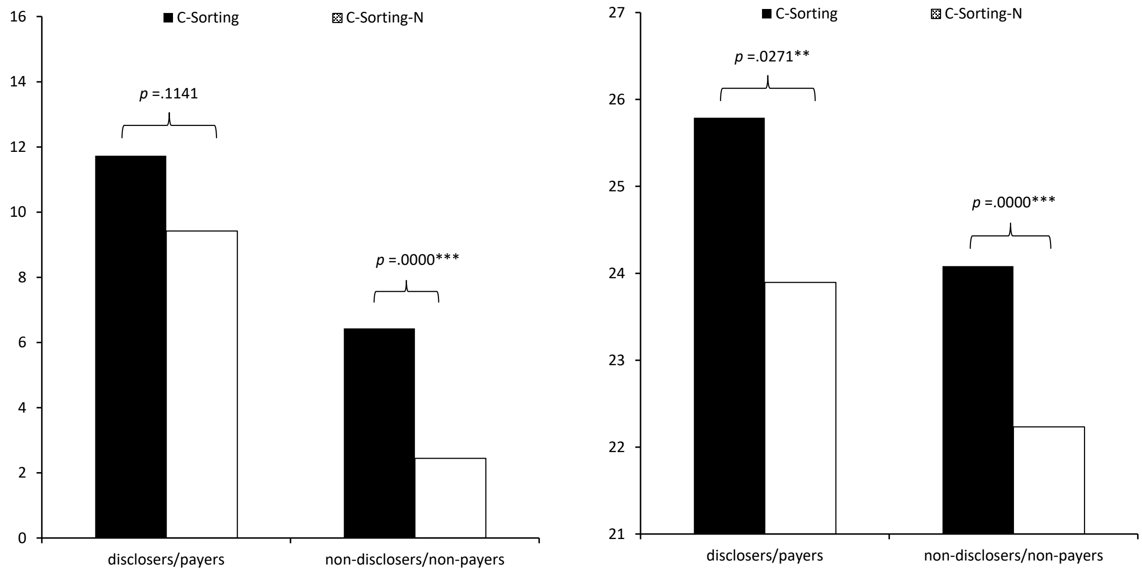
Table E.1. *The Impact of Paying a Cost without Information Disclosure in the C-Sorting-N treatment*

Dependent variable: Independent variable:	Average contributions in period t (1)	Average payoffs in period t (2)
C-Sorting dummy (#1)	5.19*** (1.40)	2.60*** (.82)
C-Sorting-N dummy (#2)	2.21 (1.58)	.80 (.93)
Constant	3.91*** (1.26)	22.3*** (.76)
# of observations	228	228
p-value (two-sided) for Wald test to the following hypothesis:		
H_0: (#1) = (#2)	.0082***	.0038***

Notes: Linear regressions. Standard errors, in parentheses, were clustered by session. Random effects were included to control for panel structure because treatment dummies are included as regressors. Dependent variable is session-average contributions or payoffs in period t. Observations of the C-Sorting, C-Sorting-N and Baseline treatments in period 2 to 20 were used in the regressions. The reference group is data from the Baseline treatment. C-Sorting dummy equals 1 for the C-Sorting treatment; and 0 otherwise. C-Sorting-N dummy equals 1 for the C-Sorting-N treatment; and 0 otherwise.

*, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

Fig. E.2. *Average Contributions and Payoffs by Disclosure Decision in the C-Sorting-N treatment*



(a) Average contribution [ECUs] (b) Average payoff [ECUs]

Notes: The session-average contributions and payoffs were first calculated by disclosure decision. Next, they were averaged by treatment. *p*-values (two-sided) were calculated based on Wald tests using the estimation of individual random effect linear regressions standard errors clustered by session. *, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.