

Department of Economics & Finance

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

Kenju Kamei

Working Paper No. 3, 2019

Department Economics and Finance
Durham University Business School
Mill Hill Lane
Durham DH1 3LB, UK
Tel: +44 (0)191 3345200
<https://www.dur.ac.uk/business/research/economics/>

© Durham University, 2019

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 studies individuals' voluntary disclosure of past behaviors and its effects in simultaneous-move dilemma interactions. Using a laboratory experiment with a finitely repeated two-player public goods game, I 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. A direct consequence of the presence of non-disclosers (and the required payment in costly disclosure treatments) is that individuals can achieve higher efficiency where disclosure is mandatory rather than voluntary. This result suggests that the elimination of an option to hide past behaviors helps enhance economic efficiency if individuals' opportunistic 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. This paper was previously circulated with a different title: "Information Disclosure and Cooperation in a Finitely-repeated Dilemma: Experimental Evidence."

1. Introduction

Voluntary disclosure of information has the potential to improve economic efficiency in sequential buyer-seller transactions under information asymmetry (see, e.g., Dranove and Jin [2010] for a survey). Three factors have been identified as possible underlying mechanisms for this effect. First, sellers' disclosure of detailed information helps buyers assess the value of products accurately (e.g., eBay Motors [Lewis, 2011]). Second, disclosure may improve the product quality delivered by the sellers (e.g., restaurants [Jin and Leslie, 2003]). Third, disclosure may improve buyer-seller matching by enabling buyers to choose which market to join (e.g., wholesale auto auctions [Tadelis and Zettelmeyer, 2015]). Past studies on information disclosure deal with cases in which one party (the seller) possesses private information on an exogenous state – the product quality, and the other party (the buyer) chooses an action after observing the seller's decision to disclose and the contents of disclosed product information. By contrast, this paper studies a case of voluntary information disclosure, focusing on the effect of disclosed information and improved matching, in a setup where (a) both parties have private information, (b) the states are *endogenous*, and (c) they make simultaneous revelation decisions.

Simultaneous-move economic transactions are increasingly common in our everyday transactions, especially online. Examples include emerging businesses designed for today's sharing economy (e.g, Uber), online-based personal relationships, and online-based professional relationships. Although scholars have intensively studied voluntary information disclosure with an exogenous one-sided state both theoretically and empirically, less attention has been paid to its effect in simultaneous-move transactions where both parties can make revelation decisions. Unlike in the sequential-move setup with information asymmetry, in a simultaneous-move interaction, an exchange between two parties take place without either side knowing the value that they will receive from the other.

The simultaneous-move interaction can be modeled using a repeated two-person prisoner's dilemma with random matching, which captures the tension between cooperation and defection.

Examples include users' interpersonal relationships in the sharing economy (e.g., conduct of Uber drivers and riders) and online-based professional relationships (e.g., between freelancers and employers). In a two-person dilemma game, the total payoff amount of two players is maximized when they both choose cooperation. However, under the assumption that players are self-interested and that everyone knows this to be the case, the Pareto-efficient outcome cannot be achieved (in a finitely repeated setting). This is because defecting on one's partner results in a higher private payoff for the defector, no matter what actions the partner takes. A source of uncertainty here stems from the private information unique to each individual.

To counter a reliance on *input* information (e.g., demographics), online platforms have developed a simpler form of behavior-based reputation mechanism built on *output* information (past interaction outcomes) as a solution to overcome cooperation problems between users. The reputation mechanism is a forced disclosure system, whereby users' ratings and reviews for past transactions are always available to potential counterparties. For example, Uber has a two-way rating system. The reputation mechanism is less costly and decentralized system, in which goods or service quality is directly judged by users themselves. Stakes involved in the transaction are small, and thus using legal channels are less beneficial for a user in case of online fraud. Prior literature has revealed that mandatory disclosure of information on individuals' past behaviors can mitigate dilemmas and achieve mutual cooperation (e.g., Engelmann and Fischbacher, 2009; Bolton *et al.*, 2004, 2005; Kamei and Putterman, 2017).¹ Nevertheless, it remains unclear how the market would look if users could *voluntarily* disclose their reputations (past behaviors).² 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

¹ The impact of exogenously given information on boosting cooperation has also been demonstrated in infinitely repeated dilemma game experiments with random matching (e.g., Camera and Casari, 2009; Kamei, 2017).

² Economists have recently started to study endogenous formation of reputational information through gossiping (e.g., Camera and Casari, 2018; Kamei and Putterman, 2018), information acquisition (e.g., Camera and Casari, 2018; Duffy *et al.*, 2013) and identity disclosure (Kamei, 2017). Camera and Casari (2018), Duffy *et al.* (2013), and Kamei (2017) used setups of infinitely repeated games, whereas Kamei and Putterman (2018) used a one-shot prisoner's dilemma game, to test the endogenous reputation formation.

wonder why mandatory, rather than voluntary, disclosure has been consistently adopted by the online platforms for users' ratings.

Although mandatory disclosure of past behaviors 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'. They are actively encouraged to reveal such scores to potential transaction counterparts even outside the platforms (e.g., Hatton 2015).³ How frequently do users voluntarily disclose their scores? What is the value of disclosure? In other situations, since individuals can usually access alternative platforms that do not have reputation mechanisms in addition to ones with mandatory disclosure (e.g., online dating services), they need to decide which platform to join. The individuals' sorting behavior is linked to their information disclosure, with implications for an effective matching process.⁴

To address these research questions, I conduct an incentivized laboratory experiment built on a finitely repeated two-person public goods game.⁵ In my design, subjects have the option to disclose their last-period contribution amounts in each period. Voluntary information disclosure has received scant attention in the literature on laboratory experiments to date. Benndorf *et al.* (2015), who recently studied voluntary information disclosure in a setup where six workers in a labor market decide whether to costly reveal their productivity, remark that: "while information disclosure has received much attention in the theoretical literature, only Forsythe *et al.* (1989) have studied unraveling in an experiment." Forsythe *et al.* investigate sellers' voluntary disclosure of the product quality in a sealed-bid auction. There are three more recent experiments that studied information unraveling in sequential seller-buyer transactions with exogenous states (Hagenbach and Perez-Richet, 2018; Jin *et al.*, 2015; Li and Schipper,

³ Another example is companies' disclosure of accounting and/or corporate information in annual reports.

⁴ Singles may disclose their detailed background information and join an online dating service where the users can see other users' background information by paying membership fees or spending time filling in registration forms and/or updating their profiles. By contrast, there exist alternative anonymous networks that do not require users to pay or spend time in submitting their detailed information.

⁵ I use a two-person public goods game, rather than a prisoner's dilemma game, because people's cooperation decisions are typically more flexible than binary choices in the prisoner's dilemma format.

2018). Unlike all these prior studies, my paper studies voluntary information disclosure of states, where the states are *endogenous* (determined by prior actions).

The closest paper to the present one is Kamei (2017), who studied people's reputation building behaviors through *identity* disclosure in an infinitely repeated prisoner's dilemma game with random matching where mutual cooperation holds as an equilibrium outcome. He found that almost everyone discloses her 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, I retain anonymity to isolate the effect of disclosed information. I show that disclosed information of own action choices per se has a positive effect. Second, I examine the effect of a disclosure rather than hiding cost, and set the cost at only 5% of the per-period endowment. This paper shows that such a small disclosure cost drastically changes the subjects' behaviors. Third, I compare reputation building behaviors between two matching protocols: random matching versus sorting. I show that individuals' cooperation behaviors (including material benefits of disclosing acts) and the impact of disclosure cost differ substantially by the matching protocol. Kamei (2017) studied people's identity disclosure behaviors only under random matching.

By using a simple disclosure format, I maintain a parsimonious design. If a player decides to disclose, 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).

The results of the experiment show that endogenously disclosed information improves cooperation under certain conditions. However, not everyone discloses, even if disclosure is free. The impact of disclosed information is not significant under random matching when disclosure involves a cost, because a large fraction of subjects are reluctant to disclose their past even if the cost is very small. 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 with sorting shows that if subjects can pay to be matched with like-minded others, disclosers are able to achieve a high degree of cooperation. The level of cooperation observed is comparable to that under mandatory information disclosure. Nevertheless, interestingly, around half of subjects choose to remain non-disclosers and keep receiving low payoffs as a result.

The rest of the paper proceeds as follows: Section 2 summarizes the experimental design, Section 3 discusses possible subject's behaviors and hypotheses of this study, Sections 4 and 5 report the experiment results and Section 6 concludes.

2. Experimental Design

This study is built on the framework of a two-person linear public good 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 in the game. Thus, the payoff of subject i ($\pi_{i,t}$) is calculated as follows:

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

Here, $c_{i,t}$ is subject i 's contribution to her joint account in period t , and subject j is i 's matched person in period t . Note that the total payoff is maximized when both subjects in a pair contribute their full endowments ($0.8 \times 2 = 1.6 > 1.0$). 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

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

This study consists of four main treatments and three control treatments. The control treatments were conducted to identify the treatment effects of voluntary information disclosure. Sections 2.1 and 2.2 describe the main and control treatments, respectively. Section 2.3 summarizes the experimental procedure.

2.1. Main Four 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 is a key difference from the prior research with information disclosure in seller-buyer transactions that deals with an exogenous one-sided state. In my setup, the state is endogenous and there is no guarantee that discloser i selects the same action as the last period.

I construct four treatments by varying two dimensions, using a 2×2 between-subjects design (Table 1). The first dimension is the size of disclosure cost: either the disclosure is free or costs one ECU. 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

both periods t and $t + 1$ because of the rich choice space, $\{0, 1, 2, \dots, 20\}$, in the contribution decisions (see Fig. C.2 for the cumulative distribution). 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 13), 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 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 two treatments.

that likewise chose to [not to] disclose in a given experimental session).⁷ I denote the four treatments as the “Costly Disclosure, Random Matching” (C-RM), “Free Disclosure, Random Matching” (F-RM), “Costly Sorting” (C-Sorting), and “Free Sorting” (F-Sorting) treatments.

Table 1. Summary of Treatments

Treatment name	Costs for disclosing	Matching protocol	The # of sessions (subjects)	Average contribution
[Main treatments:]				
C-RM	1 ECU	Random Matching: Each subject is randomly matched with another subject in a session.	4 (48)	6.02 [30.1%]
C-Sorting	1 ECU	Sorting: Each discloser (non-discloser) is matched with another discloser (non-discloser) in a session.	4 (44)	9.28 [46.4%]
F-RM	0 ECUs	Random Matching: Each subject is randomly matched with another subject in a session.	4 (44)	7.62 [38.1%]
F-Sorting	0 ECUs	Sorting: Each discloser (non-discloser) is matched with another discloser (non-discloser) in a session.	4 (44)	7.89 [39.4%]
[Control treatments:]				
Baseline	n.a.	Random Matching: Each subject is randomly matched with another subject in a session.	4 (44)	3.91 [19.6%]
C-Sorting-N	1 ECU	Sorting: Each payer (non-payer) is matched with another payer (non-payer) in a session.	4 (44)	6.31 [31.5%]
Mandatory	n.a.	Random Matching: Each subject is randomly matched with another subject in a session.	2 (24)	11.52 [57.6%]
Total			26 (292)	

Note: The numbers in squared bracket are average contributions as the percentages of endowments.

2.1.1. Conditional Contribution Schedule and Subjects’ Beliefs

I include two additional tasks. The first task is 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 .⁸ 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 my primary

⁷ If the number of disclosers (non-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% of pairings in the C-Sorting treatment and 6.9% in the F-Sorting treatment. The paper’s findings are robust regardless of whether we use only data consisting of pairs with the same preferences (two disclosers or two non-disclosers) or all the data.

⁸ In the elicitation stage, subjects are aware of their partners’ current-period disclosure decisions and the partners’ last-period contribution amounts in case that the partners selected to disclose.

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.^{9,10}

Second, I elicit subjects' cooperation types using the method of Fischbacher *et al.* (2001). Specifically, each subject is asked to answer 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 contributions (= 0, 1, 2, 3, ..., 20). Classified types based on the conditional schedules will be 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.¹¹ 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

I additionally conducted three control treatments (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 will serve 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

⁹ In the experiment, these elicitation tasks were only included 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.

¹⁰ In a finitely repeated linear public goods game, Gächter and Renner (2010) showed that incentivized elicitation improves belief accuracy only a little, while it significantly influences subjects' decisions to contribute. 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 complexity of design and because my focus is on subjects' actual disclosure and contribution behaviors, I did not employ this method.

¹¹ In this task, subjects are randomly assigned into groups of four and make two kinds of contribution decisions in a linear public goods game with a MPCR of 0.4. The first decision is the conditional contribution decisions 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.

efficiency among the four main treatments (Section 4). At the onset of the Mandatory treatment, subjects are randomly assigned to a community of six members and the group composition stays the same during the experiment.¹² In each period, subjects are randomly matched with 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 and 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 information is not disclosed. Subjects experience two stages in all rounds after period 1. In the first stage, they decide whether or not to pay one ECU. Subjects who pay (do not pay) one ECU are randomly matched with another who pays (does not pay) it. I call a player who pays (does not pay) the fee the “payer” (“non-payer”). However, their last-period contributions are *not* revealed. The data of this treatment will be used to study the relative importance of (a) the presence of a matching cost and (b) disclosed information, for determining the high performance in the C-Sorting treatment. Since this treatment is not directly related to the hypotheses of the paper, I relegate the details of the results to Appendix D.

2.3. Experimental Procedure

A total of 26 sessions (16 for the main four 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 are exchanged for £1) at the end of the experiment.¹³ The average payment (including £3 show-up fee) was £15.61. The average

¹² 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 5.5).

¹³ 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.

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 if we assume that they believe their peers would always select defection. 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 treatments as these actions can be taken for free and the peers would always select defection; hence, players would randomly decide whether to disclose.

Hypothesis 1: *Standard Theory Prediction.*

(a) No one costly discloses their state (last-period contribution behavior) in the C-RM or C-Sorting treatment. (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.

On the other hand, theoretically 'rational' cooperation is possible even in a finitely repeated dilemma game if we assume that some people believe that some of their peers act based on a discriminating strategy, such as tit-for-tat (e.g., Kreps *et al.*, 1982). Nevertheless, in the absence of institutions to assist people's cooperation behaviors, a rich 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 public goods games, they decrease their levels of cooperation steadily over time (Ledyard, 1995; Chaudhuri, 2011).¹⁴

With voluntary information disclosure, cooperation could instead evolve, similar to the logic of Kreps *et al.* (1982). In my framework, a typical example of such a discriminating

¹⁴ Similar free-riding dynamics is reported when the group size is two, irrespective of whether partner or random matching is used (e.g., Kamei, 2019).

strategy is the ‘conditional cooperation strategy’ (e.g., Fischbacher *et al.*, 2001).¹⁵ The prevalence of conditional cooperators (i.e., players who follow the conditional cooperation strategy) is well-documented. In the theoretical analysis, I will assume, for simplicity, that a player 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.¹⁶ The steady decline of contributions in a 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).

Let us discuss the possibility of rational cooperation by examining the behavior of a material payoff maximizer i , assuming that there exist players who choose to disclose and then act according to the conditional cooperation strategy, in addition to free rider types (players who contribute zeros no matter what circumstance). I use $p_{cc} \times 100\%$ and $p_{FR} \times 100\%$ to refer to the percentages of the conditional cooperators and the free riders, respectively ($p_{cc} + p_{FR} = 1$). I assume that the free riders do not disclose the states when disclosure is costly, as described by Hypothesis 1(a). When disclosure is free, I assume that free riders disclose them with a probability of 50%, as described by Hypothesis 1(b). Having unmasked free riders would not affect our calculation results under random matching, because of our assumption that conditional cooperators contribute zero towards disclosers whose states are zero, but it would decrease the value of information disclosure under sorting. For simplicity, I set $x = 1$ (a perfect conditional cooperator).¹⁷ Lastly, suppose that i correctly anticipates her peers’ behaviors and that her belief on p_{cc} is correct.

¹⁵ 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).

¹⁶ 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 my setup is someone who contributes nothing.

¹⁷ The basic implication we have for a material payoff maximizer’s behaviors does not change even if we set $x < 1$. In this theoretical analysis, I also assume that each discloser (non-discloser) is randomly matched with another discloser

In this illustrative framework, it would be materially beneficial for material payoff maximizer i to mimic the behavior of a conditional cooperator if there is a sufficiently large fraction of conditional cooperators in the community where disclosers operate. The detail is provided in Appendix B. This prediction can be derived by setting up a Hamiltonian and applying the Maximum principle (e.g., Sethi and Thompson, 2006).

Proposition 1: *If the percentage of conditional cooperators in the community where disclosers operate (p) is sufficiently large that $p > 0.25$ in a given treatment, then: in all periods before period 20, i selects cooperation and then discloses the state; in the end period (period 20), i discloses the state and then selects $c_{i,20} = 0$. If $p < 0.25$, the prediction is the same as Hypothesis 1.*

First, let us assume that no free riders other than i mimic the behavior of a conditional cooperator. Then, while $p = p_{cc}$ in the C-RM and F-RM treatments, $p = 1$ and $p = p_{cc}/[p_{cc} + 0.5 \cdot p_{FR}] \geq p_{cc}$ in the C-Sorting and F-Sorting treatments, respectively, because of the sorting mechanism.¹⁸ These values for p and Proposition 1 suggest three behavioral patterns. First, i 's incentives to build a cooperative reputation would be stronger with than without sorting. Second, the disclosure cost does not affect the mimicking incentives under random matching, as the cost is too small, just one ECU (see also Appendix B). Third, a positive disclosure cost does raise the mimicking incentives under the sorting condition, because the cost deters some opportunistic free riders from invading into the set of disclosers.¹⁹ Proposition 1 also suggests the same behaviors between the two random matching treatments and the Mandatory treatment, because of our assumption that no free riders other than i would mimic the behavior of conditional cooperator (i.e., $p = p_{cc}$ in the Mandatory treatment).

Hypothesis 2: *Incentives to Mimic the Behavior of a Conditional Cooperator*

(non-discloser) always under the sorting condition for simplicity. It is worth noting that the exception remarked in footnote 7 would decrease the mimicking incentives in the C-Sorting and F-Sorting treatments. It would not, however, change the comparison of subjects' behaviors between the two matching protocols.

¹⁸ $p_{cc}/[p_{cc} + 0.5 \cdot p_{FR}] - p_{cc} = p_{cc} \cdot (1/[p_{cc} + 0.5 \cdot p_{FR}] - 1) \geq 0$.

¹⁹ This analysis implies that p and thus the disclosure rate could be higher in the C-Sorting than in the F-Sorting treatment.

(a) *Sorting improves cooperation: $D(\text{C-Sorting}) > D(\text{C-RM})$ and $D(\text{F-Sorting}) > D(\text{F-RM})$.*

Likewise, $C(\text{C-Sorting}) > C(\text{C-RM})$ and $C(\text{F-Sorting}) > C(\text{F-RM})$.

(b) *$D(\text{F-RM}) \geq D(\text{C-RM})$. However, $C(\text{F-RM}) = C(\text{C-RM})$.*

(c) *The positive disclosure cost improves cooperation under sorting: $C(\text{C-Sorting}) > C(\text{F-Sorting})$.*

(d) *$C(\text{F-RM}) = C(\text{Mandatory})$.*

Here, $D(\cdot)$ and $C(\cdot)$ refer to the disclosure rate and average contribution amount, respectively, in a given treatment. Proposition 1 states that a sufficiently large fraction of conditional cooperators must be present to induce material payoff maximizers to select cooperation. In an extreme situation, players do not need to exhibit non-standard preferences, as discussed in Kreps *et al.* (1982). If a sufficiently large number of players believe that a large percentage of the peers will disclose and choose (mimic) such a conditional cooperation strategy, beliefs can become self-fulfilling and cooperation may evolve even in the presence of exclusively selfish types.²⁰

If we instead assume that some free riders other than i mimic conditional cooperators, then i 's mimicking incentives change in some conditions. While Hypotheses 2(a)-(c) do not change with the prevalence of the mimickers, Hypothesis 2(d) does. To see this, suppose that $p_m \times 100\%$ of free rider types mimic conditional cooperators. This change in the assumption alters p (the likelihood that a material payoff maximizer i is matched with a person who acts according to the conditional cooperation strategy) as follows:

$$\left\{ \begin{array}{l} p^{RM} = p_{cc} + p_{FR} \cdot p_m^{RM} \text{ for both the C-RM and F-RM treatments.} \\ p^S = 1 \text{ for the C-Sorting treatment; } p^S = \frac{p_{cc} + p_{FR} \cdot p_m^S}{p_{cc} + p_{FR} \cdot p_m^S + 0.5 \cdot p_{FR} \cdot (1 - p_m^S)} \text{ for the F-Sorting treatment.} \\ p^M = p_{cc} + p_{FR} \cdot p_m^M \text{ for the Mandatory treatment.} \end{array} \right.$$

²⁰ While these theoretical discussion suggest the possibility of information unraveling, subjects may fail to do so. Although the setups are different, there are several past unraveling experiments with sequential transactions where one party discloses its quality (e.g., product quality). However, there are no clear agreements as to when information unraveling happens: while Forsythe *et al.* (1989) and King and Wallin (1991) showed information unravelling happens in auctions, Benndorf *et al.* (2015), Li and Schipper (2018) and Jin *et al.* (2015) explained that information revelation can be less common than theory predicts in other setups.

Here, the superscripts, RM , S and M , refer to the random matching, sorting, and Mandatory conditions, respectively. We can reasonably assume that p_m is the same for the C-RM and F-RM treatments since the size of disclosure cost does not affect i 's mimicking incentives (see Proposition 1 and Appendix B). Since endogenous regrouping is known to encourage selfish types to mimic cooperative types, we can also assume that $p_m^S > p_m^{RM}$. This latter assumption would not affect the sign of the comparison between the random matching and sorting treatments. Thus, it is clear that Hypotheses 2(a)-(c) hold for the revised p^{RM} and p^S .

In contrast, Hypothesis 2(d) does not hold because we can assume that $p_m^M > p_m^{RM}$ (and so $p^M > p^{RM}$). Kamei (2017) showed that: (i) given an option to hide IDs, a non-negligible fraction of individuals do not disclose their IDs and select defection, 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 behaviors might plausibly 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 of conditional cooperators. The condition $p^M > p^{RM}$ suggests that $C(Mandatory) > C(C-RM) = C(F-RM)$. On the other hand, no clear prediction is possible for a comparison between the F-Sorting versus Mandatory treatment. p_m^S may be larger than p_m^M considering that sorting assures a better matching among like-minded individuals. However, the opposite ($p_m^M > p_m^S$) may also be possible if eliminating the option to hide encourages those who would hide and then defect in the F-Sorting treatment to strategically build a cooperative reputation under mandatory disclosure.

Hypothesis 3: *Mimicking Incentives in the Presence of Mimickers of Conditional Cooperators.* $C(C-Sorting) > C(Mandatory) > C(C-RM) = C(F-RM)$. No clear prediction is possible between the F-Sorting and Mandatory treatments. Hypotheses 2(a)-(c) still hold.

4. Treatment Effects of Disclosure and Better Matching

I will give an overview of subjects' decisions in Section 4.1. I then formally study the treatment differences of the subjects' behaviors in Section 4.2 and material benefits of disclosing acts in Section 4.3.

4.1. Overview of the Experiment

I first examine the effect of exogenously disclosed information using the data of the Baseline and Mandatory treatments. Fig. 1 shows subjects' average behaviors while dividing the data into the first half (periods 2-10) and the second half (periods 11-20) of the experiment.^{21,22} It suggests that communities can achieve high efficiency if everyone is forced to disclose their state $S_{i,t}$ (panels (a) and (c)). The overall average contribution in the Mandatory treatment was 11.53 ECUs – about three times as large as that in the Baseline treatment (3.83 ECUs). A difference was also observed for subjects' beliefs (panel (b)). The average beliefs in the Mandatory and Baseline treatments were 11.45 and 3.99 ECUs, respectively. The strong effect of mandatory disclosure on improving cooperation is consistent with the prior studies that found positive effects in the contexts of helping game (e.g., Bolton *et al.*, 2005) and investment game (e.g., Bolton *et al.*, 2004).

Fig. 1 further includes the data of the main four treatments. While the majority of hypotheses 2 and 3 hold, the theoretical analyses do 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)).²³ Consistent with Hypothesis 2(a), the disclosure rate

²¹ Data in period 1 was excluded in comparing the subjects' behaviors across the treatments, because they were not given an opportunity to disclose the state in that period in the treatments with information disclosure. Results are nevertheless similar if data in period 1 is included (the results are omitted to conserve space).

²² 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 we use for comparison. See Appendix Fig. C.1 for the trends of subjects' period-by-period decisions.

²³ As I assumed in the theoretical analysis in Section 3, panel (d) also indicates that subjects correctly anticipated the peers' disclosure behaviors: the subjects' average beliefs on the peers' disclosure rates transited almost parallel to the realized percentages of disclosers in the C-RM and F-RM treatments.

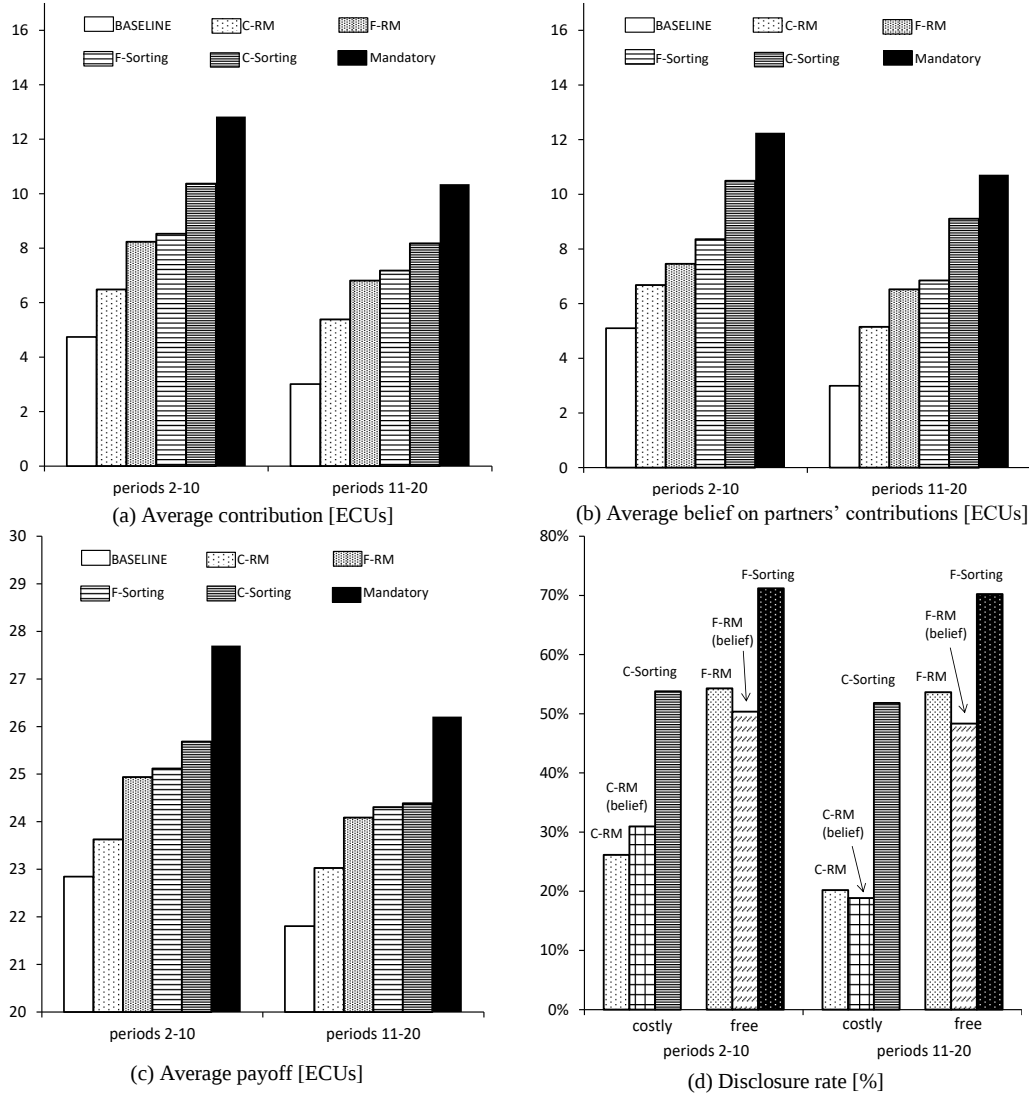
(the percentage of the cases in which subjects chose to disclose) was higher with sorting than under random matching, although it remained well below 100% even with sorting. The session-average disclosure rates in the C-Sorting and F-Sorting treatments were 52.8% and 70.7 %, respectively. Consistent also with Hypothesis 2(b), the presence of a disclosure cost undermined subjects' disclosure rates under random matching. However, a similar impact of disclosure cost was also seen with sorting, at odds with the suggestion in Section 3 that a positive disclosure cost might increase the number of mimickers by effectively discouraging opportunistic free riders from invading into the set of disclosers. It may be that some subjects did not recognize the positive effect of the disclosure cost.

Second, voluntary information disclosure has a positive effect on improving cooperation, and accordingly payoff, in all the four treatments (see 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 (free or costly). However, the difference between the F-RM than F-Sorting treatments was small.

Third, the positive disclosure cost helped improve cooperation with sorting, consistent with Hypothesis 2(c). This implies that a higher disclosure rate due to free disclosure is not helpful in improving contribution. By clear contrast, the positive disclosure cost strongly undermined cooperation under random matching, unlike Hypothesis 2(b). The cost may have driven some cooperative types to non-disclosure and conservative behavior in the dilemma interactions. This interpretation is consistent with the low disclosure rate (only 23.0% on average) in the C-RM treatment, which diminishes material payoff maximizers' incentives to mimic the behavior of a conditional cooperator. The result implies that subjects have a clear discontinuity in their reputation building behaviors between positive and zero costs (e.g., Shampanier *et al.*, 2007).

Fourth, the efficiency under the mandatory disclosure was the highest among all the treatments. This is not consistent with Hypothesis 2 or 3.

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



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, minus average disclosure costs. See Appendix Fig. C.1 for period-by-period average behaviors.

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 perhaps occurred because disclosure involved a cost and no sorting mechanism was present. Thus, subjects, even conditional cooperators, might have gradually perceived that disclosure was not worth the cost.

4.2. Disclosure Behaviors and the Impact of Disclosed Information

For a formal analysis of the treatment differences, I estimate a regression model (Table 2), in which the dependent variable is session-average disclosure rates (columns (1) and (2)), session-average contribution amounts (columns (3) and (4)) or session-average payoffs (columns (5) and (6)) in each period. The independent variables include treatment dummies.

I first confirm that subjects' disclosure rates were significantly positive in all four treatments (column (1)). Nevertheless, 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. In addition, a 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 each of the four treatments with voluntary information disclosure.* (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 estimation also shows that the impact of information disclosure on efficiency depends on the disclosure rate. The impact is not significant in the C-RM treatment, whose disclosure rate was the lowest in the four treatments (columns (3) and (5)). In contrast, it is significant in the F-RM treatment. These findings confirm the earlier discussions that subjects' behaviors were sensitive to the presence of disclosure cost under random matching, unlike Hypothesis 2.

The sorting mechanism, however, did improve cooperation substantially when disclosure was costly. The average contribution and payoff were both significantly larger in the C-Sorting than in the C-RM treatment. By contrast, the sorting mechanism had little impact on efficiency when disclosure was free. Overall, columns (3) and (5) of Table 2 indicate that there were no

treatment differences in the level of cooperation or payoff between the three treatments (F-RM, C-Sorting and F-Sorting), in which the disclosure rates were more than 50%.

This regression analysis also confirms the outperformance of mandatory information disclosure, not predicted by Hypothesis 2 or 3. The efficiency (whether in contributions or

Table 2. *Treatment Differences in Subjects' Behaviors*

Dependent variable:	Disclosure rates		Average contributions		Average payoffs	
Independent variable:	(1)	(2)	(3)	(4)	(5)	(6)
(#1) C-RM dummy {= 1 for the C-RM treatment; 0 otherwise}	---	---	1.99 (1.39)	1.91 (1.20)	.97 (.82)	.83 (.71)
(#2) C-Sorting dummy {= 1 for the C-Sorting treatment; 0 otherwise}	.29*** (.07)	.24** (.10)	5.19*** (1.38)	5.57*** (1.52)	2.60*** (.80)	2.79*** (.86)
(#3) F-RM dummy {= 1 for the F-RM treatment; 0 otherwise}	.30*** (.06)	.25*** (.043)	3.57** (1.52)	3.74*** (1.12)	2.14** (.91)	2.25*** (.67)
(#4) F-Sorting dummy {= 1 for the F-Sorting treatment; 0 otherwise}	.48*** (.05)	.37*** (.062)	4.07*** (1.58)	3.67*** (1.34)	2.44*** (.94)	2.20*** (.81)
(#5) Mandatory dummy {= 1 for the Mandatory treatment; 0 otherwise}	---	---	7.62*** (1.24)	8.58*** (2.04)	4.57*** (1.29)	5.15*** (1.23)
Period Number {= 2, 3, ..., 19, 20}	---	-.01*** (.002)	---	-.16*** (.039)	---	-.097*** (.023)
Constant	.23*** (.041)	.31*** (.03)	3.91*** (1.24)	5.70*** (1.06)	22.3*** (.74)	23.4*** (.64)
Number of observations	304	304	456	456	456	456
Reference group	C-RM	C-RM	Baseline	Baseline	Baseline	Baseline
<i>p</i> -value (two-sided) for Wald χ^2 tests to the following hypothesis:						
H ₀ : (#1) = (#2)	---	---	.0003***	.0028***	.0004***	.0027***
H ₀ : (#1) = (#3)	---	---	.1450	.0059***	.0612*	.0002***
H ₀ : (#1) = (#4)	---	---	.0729*	.0760*	.0290**	.0193**
H ₀ : (#1) = (#5)	---	---	.0027***	.0003***	.0012***	.0001***
H ₀ : (#2) = (#3)	.8245	.9270	.1266	.1131	.4526	.3748
H ₀ : (#2) = (#4)	.0030***	.2305	.3273	.1649	.8151	.4365
H ₀ : (#2) = (#5)	---	---	.1937	.1448	.0735*	.0487**
H ₀ : (#3) = (#4)	.0001***	.0549*	.6995	.9355	.6995	.9355
H ₀ : (#3) = (#5)	---	---	.0398**	.0068***	.0398**	.0068***
H ₀ : (#4) = (#5)	---	---	.0784*	.0111**	.0784*	.0111**

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 was used. Observations from period 2 to 20 in the four treatments with information disclosure [in the four treatments with information disclosure, the Baseline treatment and the Mandatory treatment] were used as data in columns (1) and (2) [in columns (3) to (6)]. In addition to the Period Number variable, the interaction term between each treatment dummy and the Period Number variable was added as control variables in columns (2), (4), and (6); because all the interaction terms, except three cases, fail to obtain significant coefficients, I did not include them in this table to conserve space. The three exceptions are (i) the interaction term between variable (#4) and the Period Number variable in column (2), which has a significantly positive coefficient, while the coefficient estimate for variable (#4) becomes smaller compared with column (1); and (ii) the interactions terms between variable (#5) and the Period Number variable in column (4) (column (6)), which has a significantly negative coefficient, while the coefficient estimate for variable (#5) becomes larger compared with column (3) (column (5)).

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

payoffs) in the Mandatory treatment is significantly higher than in the C-RM and F-RM treatments. It is also weakly significantly higher than in the F-Sorting treatment. Lastly, although the average contribution in the Mandatory treatment is not significantly different from that in the C-Sorting treatment, the average payoff is significantly higher in the Mandatory than in the C-Sorting treatment at the 10% level, because of some subjects' costs required for disclosure and the presence of the inefficient non-disclosers in the latter treatment.

Result 2: (i) *Costly disclosure enhanced neither cooperation nor payoffs significantly in the C-RM treatment.* (ii) *The average contribution and payoff were both significantly higher in the C-Sorting than in the C-RM treatment.* (iii) *The average contributions and payoffs were at similarly high levels for the C-Sorting, F-RM and F-Sorting treatments.* (iv) *The average contributions and payoffs were both significantly (weakly significantly) higher in the Mandatory treatment than in the C-RM and F-RM treatments (the F-Sorting treatment). The average payoff in the Mandatory treatment was weakly significantly higher, compared with the C-Sorting treatment.*

4.3. Subjects' Benefits of Disclosing Acts

I discussed in Section 3 that material payoff maximizers have material incentives to contribute large amounts and then disclose them under certain conditions. Consistent with this discussion, disclosers contributed more strongly than non-disclosers in all four treatments (Fig. 2(a)). There is a weakly significant difference in the disclosers' behaviors between the C-Sorting and F-Sorting treatments, similar to the treatment difference between the two treatments discussed in Section 4.1. Aside from this, no other treatment differences are found (Appendix Table C.1).

One might expect from the disclosers' strong contribution behaviors that subjects had strategic reputation building incentives in all four treatments. This is not the case. As shown in Fig. 2(b), there is a clear contrast regarding 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.²⁴ This suggests that with 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 *higher* payoffs in expectation than non-disclosers.²⁵ This underscores the role of sorting in making information disclosure materially beneficial.

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.*²⁶

As I explained in Section 2, the Mandatory treatment was constructed for a comparison with the C-Sorting treatment. The (average) sub-community size within a session is six in both of the treatments. 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 cooperate if they are all forced to disclose, and hence that the mandatory disclosure system functions more effectively than the voluntary disclosure system, since it eliminates an option to hide their past behaviors.

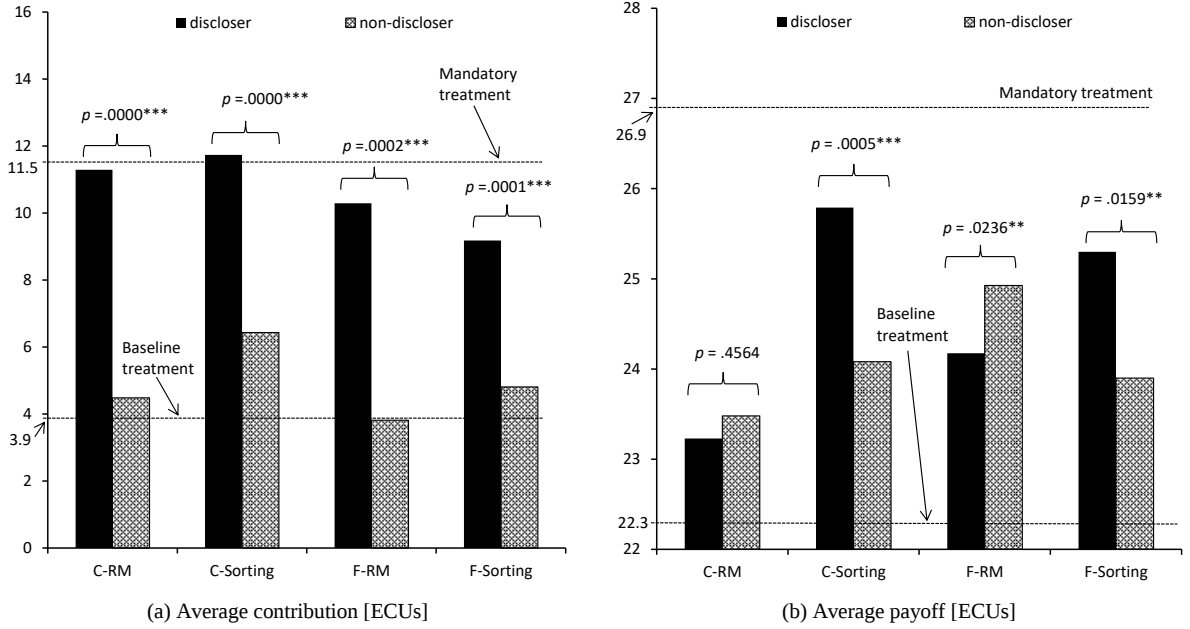
The theoretical discussions in Section 3 explain Result 3(ii) well. I discussed that a subject's strategic reputation building incentive is increasing in p (the likelihood to be matched

²⁴ The lack of significance in the C-RM treatment is due to the non-disclosers' low average payoff in the C-RM treatment (Fig. 2(b)). The non-disclosers were less frequently matched with disclosers in the C-RM than in the F-RM treatment, as the number of disclosers was much less in the former than in the latter treatment.

²⁵ 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 the peers than those in the F-Sorting treatment, the former needed to incur a cost for disclosure.

²⁶ As an anonymous referee pointed out, 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 prisoner's dilemma game). Kamei and Putterman (2017), using a repeated supergame design with or without history information, 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 when they repeat the supergame.

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



Notes: Each bar was calculated in the following two steps. I first calculated session-average contributions (payoffs) by disclosure decision. I next averaged them by treatment. Each p -value (two-sided) in Fig. 2 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 standard errors clustered by session (as in Table 2). The average contributions (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.

with someone who acts according to the conditional cooperation strategy [i.e., conditional cooperator or mimicker]). I also discussed that p would be higher with sorting than under random matching and that p would be higher in the C-Sorting than in the F-Sorting treatment. The behavioral patterns outlined above bear this out.

5. Subjects' Disclosure Decisions and Action Choices

The theoretical analyses in Section 3 did not perfectly explain the subjects' behaviors (Section 4). This section is devoted to a detailed analysis of the individual subjects' behaviors.

5.1. Conditional Disclosure Behaviors under Random Matching

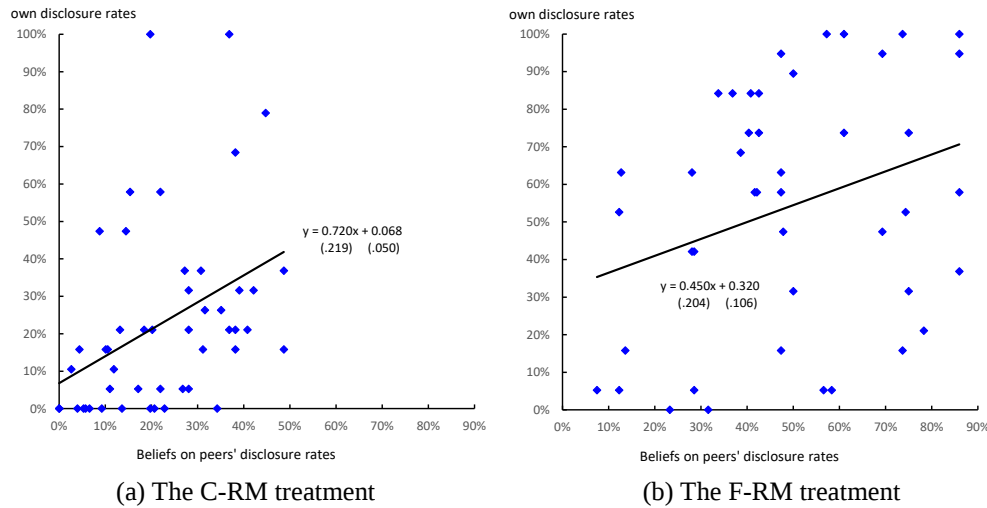
Analyses in Section 4 largely confirmed the theoretical implication in Section 3 that a material payoff maximizer would mimic the behavior of conditional cooperator, provided that p is sufficiently high (e.g., Result 1(ii), Result 3(ii)). The importance of p can be checked further

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) should depend on their beliefs on p .

Consistent with the theoretical implication, subjects' average disclosure behaviors were significantly positively correlated with their beliefs regarding the peers' disclosure rates (Fig. 3). Thus, the model of subjects' conditional cooperation strategies can help explain our 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.²⁷

Result 4: *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. 3: *Subjects' Disclosure Rates and Beliefs on the Peers' Disclosure Behaviors*



Notes: Each point indicates each subject's data. 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.

5.2. Subjects' Disclosure Decisions and Their States

While the theoretical analysis in Section 3 suggests that material payoff maximizers have

²⁷ 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).

incentives to pretend to be a conditional cooperator under mild conditions in all the main treatments, a non-negligible fraction of subjects did not disclose their states $S_{i,t}$. This was the case even with sorting, despite Result 3(ii). In order to examine this behavioral pattern in greater depth, I checked the relationship between subjects' states and their disclosure decisions (Fig. 4). Three clear patterns emerged.

First, a subject's likelihood to disclose her reputation state $S_{i,t}$ 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 subjects with high (low) $S_{i,t}$ maintained high (low) status over time, even though they could change their reputation status easily in the experiment (recall that $S_{i,t}$ is only dependent on her last-period contribution amount). In Section 3, I assumed that conditional cooperators would contribute nothing towards non-disclosers. This assumption means that all subjects with $S_{i,t} > 0$ would choose to disclose when disclosure is free. This does not perfectly explain the subjects' behaviors since for example, around 76.8% and 43.3% of subjects in the F-RM treatment hid their states when $S_{i,t} \in [1, 4]$ and $S_{i,t} \in [5, 8]$, respectively (see panel (b)).

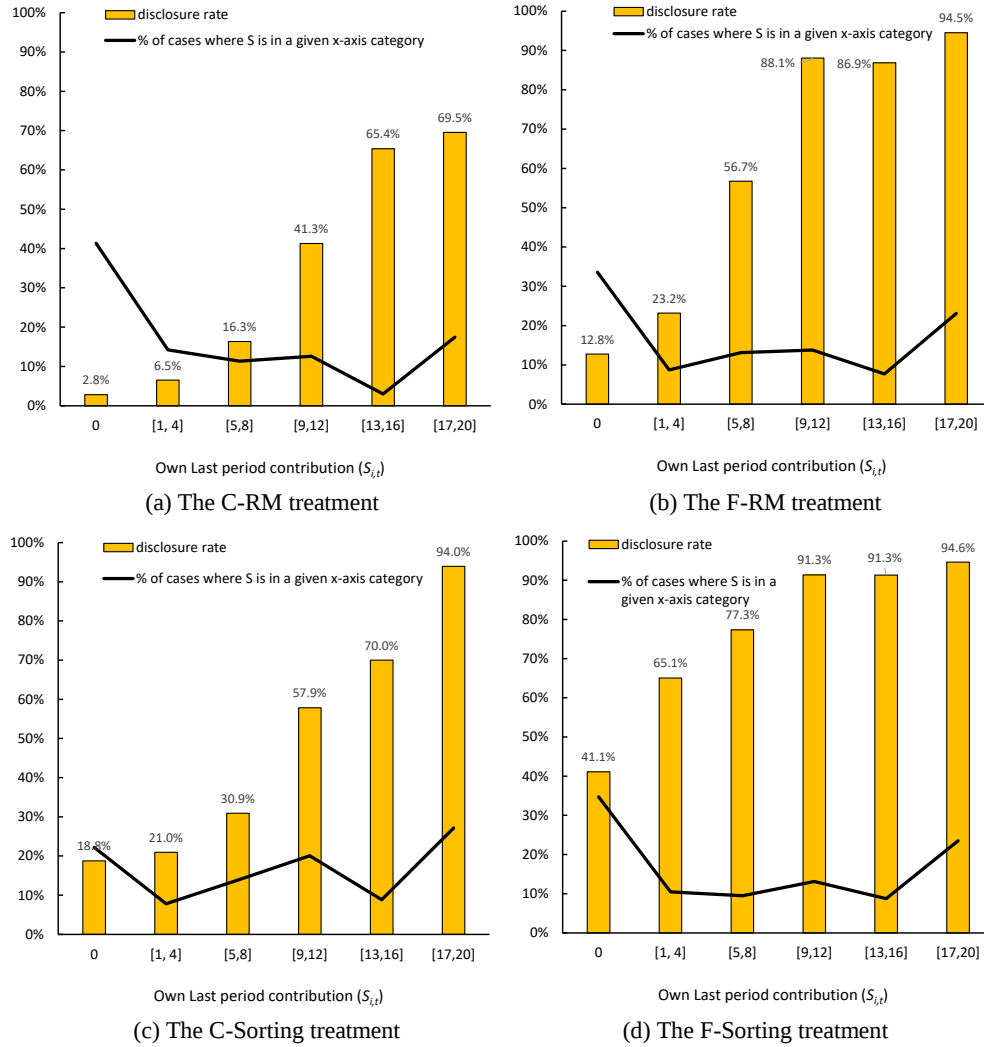
Second, we see that consistent with Hypothesis 2(a), sorting nevertheless raised subjects' disclosure rates for each category of $S_{i,t}$ except the highest. If $S_{i,t} \in [17, 20]$, the average disclosure rate was 94.5% even in the F-RM treatment.

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

Result 5: (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 that of positive disclosure cost occurred regardless of the size of $S_{i,t}$.*

Fig. 4 also demonstrates that for almost all categories of $S_{i,t}$, subjects' disclosure behaviors were stronger in the F-RM than in the C-Sorting treatment. Result 3(ii) can thus be explained by the difference in the matching protocol. However, it is possible that subjects might

Fig. 4: Subjects' Disclosure Rates by the Size of their State



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

have also behaved more strategically under random matching than with sorting. To study the validity of this possibility, I conducted a regression analysis in which the dependent variable is a dummy that equals 1(0) if a subject chose to (not to) disclose her state in period t . As shown in Appendix Table C.2, disclosers in a sorting treatment in period t are significantly more likely than disclosers in the corresponding random matching treatment to disclose their states again in period $t + 1$ (see the estimates of the constant terms). This implies that some opportunistic individuals may switch between disclosing and hiding to exploit cooperative partners under random matching.

Result 6: *Subjects switched between disclosing and hiding the states more frequently in the random matching treatments than in the sorting treatments.*²⁸

5.3. Subjects' Responses to Peers' Disclosure Decisions

Some subjects' hiding behaviors seen in Fig. 4 and Result 5 may be explained by subjects' responses to non-disclosers or to states of disclosers. 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 such hiding behaviors materially beneficial under random matching as seen in Result 3(ii). Yet understanding our 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 3(a)).^{29,30} 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 3(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 3(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 3 suggest that the material benefit from disclosure exceeds the cost of one ECU for $S_{j,t} > 7.27$.³¹ In fact, subjects' disclosure rates were low for all $S \leq 8$ (Fig. 4(a)).

By contrast, hiding behaviors for subjects with $S_{i,t} > 0$ in the F-RM treatment cannot be

²⁸ Some switching behaviors were seen in all the four treatments with voluntary information disclosure. First, the lower the amounts that disclosers contributed in period $t - 1$, the more likely they were to become non-disclosers in period t in all four treatments. Similarly, the higher the amounts that non-disclosers contributed in period $t - 1$, the more likely they were to become disclosers in period t in the C-RM, F-RM and F-Sorting treatments.

²⁹ I also ran the same regressions by 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) has a significantly positive coefficient in every column. The results are omitted to conserve space.

³⁰ Observations in period 20 were not included in the regressions since the usual end-game defection was observed. Results are similar even if observations in period 20 are included (the results are omitted to conserve space).

³¹ If we use the parameters: .43 in panel (a) and .32 in panel (b1) 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, we can calculate an expectation of i 's additional contribution as $.32 \times .43 \times S_{j,t}$ if $S_{j,t}$ is observable.

Table 3. 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 $2 \leq t \leq 19$.

Independent variable:	Matching:	Random matching				Sorting			
	Decision-maker i :	Discloser		Non-discloser		Discloser ^{#1}		Non-discloser ^{#2}	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		C-RM	F-RM	C-RM	F-RM	C-Sorting	F-Sorting	C-Sorting	F-Sorting
(i) i 's belief on partner j 's contribution in period t ($C_{i,j,t}^b$)		.43** (.083)	.50*** (.076)	.65*** (.061)	.44** (.11)	.45*** (.046)	.55*** (.078)	.39*** (.045)	.26 (.18)
Constant		7.65*** (.66)	6.42*** (.65)	1.02** (.32)	1.69* (.59)	6.77*** (.56)	4.03*** (.69)	3.58*** (.32)	4.36*** (.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 $2 \leq t \leq 19$ ($C_{i,j,t}^b$).^{#3}

Independent variable:	Decision-maker i :	Discloser		Non-discloser	
		(1) C-RM	(2) F-RM	(3) C-RM	(4) F-RM
Partner-disclose dummy {=1 if i 's period t partner j disclosed his state in period t }		1.66 (2.38)	-.56 (.78)	-.55 (.72)	-.22 (.56)
Partner-disclose dummy $\times j$'s state [i.e., period $t - 1$ contribution amount]		.33** (.094)	.46** (.10)	.32*** (.041)	.34** (.098)
Constant		6.50*** (.26)	5.76*** (.38)	4.33*** (.041)	3.10*** (.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}

Independent variable:	C-Sorting	F-Sorting
	(1)	(2)
Subject i 's period t matched discloser j 's state [i.e., period $t - 1$ contribution amount]	.48*** (.053)	.42*** (.031)
Constant	6.03*** (.74)	4.51*** (.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 we instead use subject random effect ordered probit regressions (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 we exclude these observations. ^{#2} A small number of the non-disclosers were matched with disclosers. Results in columns (7) and (8) are similar even if we exclude these observations. ^{#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.

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 3(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 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.³² 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 $S_{i,t} = 0$ (note that $(9.48 - 6.15) \times 0.8 > 1$). Likewise, disclosure was a materially beneficial action in the F-Sorting treatment because non-disclosers' average contribution was 5.50, and $6.51 + .231 \times S_{i,t} > 5.50$.

Result 7: (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.*

In summary, analyses in Sections 5.2 and 5.3 suggest that the failure of information unraveling is due to the weak disclosure behaviors taken by those with low $S_{i,t}$. But then, why did they not attempt to build cooperative reputations using the information disclosure system? Perhaps non-disclosers mistakenly believed that their matched partners would not be skeptical

³² 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 3.

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).

5.4. Cooperation Types and Strategic Reputation Building Behaviors

Let us now turn our attention to how subjects' disclosure and contribution behaviors differ by cooperation type. As detailed below, the data reveal that (a) some selfish subjects strategically contributed positive amounts for future disclosure in line with the idea that they mimic the behaviors of conditional cooperators, but (b) not all conditional cooperators behave cooperatively.

The conditional contribution schedule (Fischbacher *et al.*, 2001) elicited from each subject can be used for this analysis.³³ I focus on two types: "conditional cooperators" and "free riders."³⁴ Similar to Fischbacher *et al.*, I define those whose own contributions and the others' average contributions are significantly positively correlated at least at the 5% level (according to spearman's ρ) as the conditional cooperators. I define those whose own contributions are always zero in the task as free riders.³⁵ The percentages of conditional cooperators and free riders who disclosed the 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 his or her state.

The fit of the Fischbacher *et al.*'s (2001) method can be checked by looking at subjects' contribution amounts in period 20. A subject would contribute nothing in period 20 if he or 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. 5). This implies that the classified types are good indicators to measure subjects' contribution behaviors.

³³ The average conditional contribution schedule exhibits that of a conditional cooperator (Appendix Fig. C.3).

³⁴ Fischbacher *et al.* (2001) found a type called "hump-shared contributor." Some subjects in this study also exhibited this pattern. But only around 8.3% of subjects were classified as hump-shaped in the task.

³⁵ 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.

The data reveal that not only conditional cooperators but also free riders contributed similarly large amounts in each treatment before period 20 (each panel (a) of Fig. 5). This clearly suggests that some free riders mimicked the behavior of conditional cooperator by strategically contributing large amounts and disclosing them in following periods.

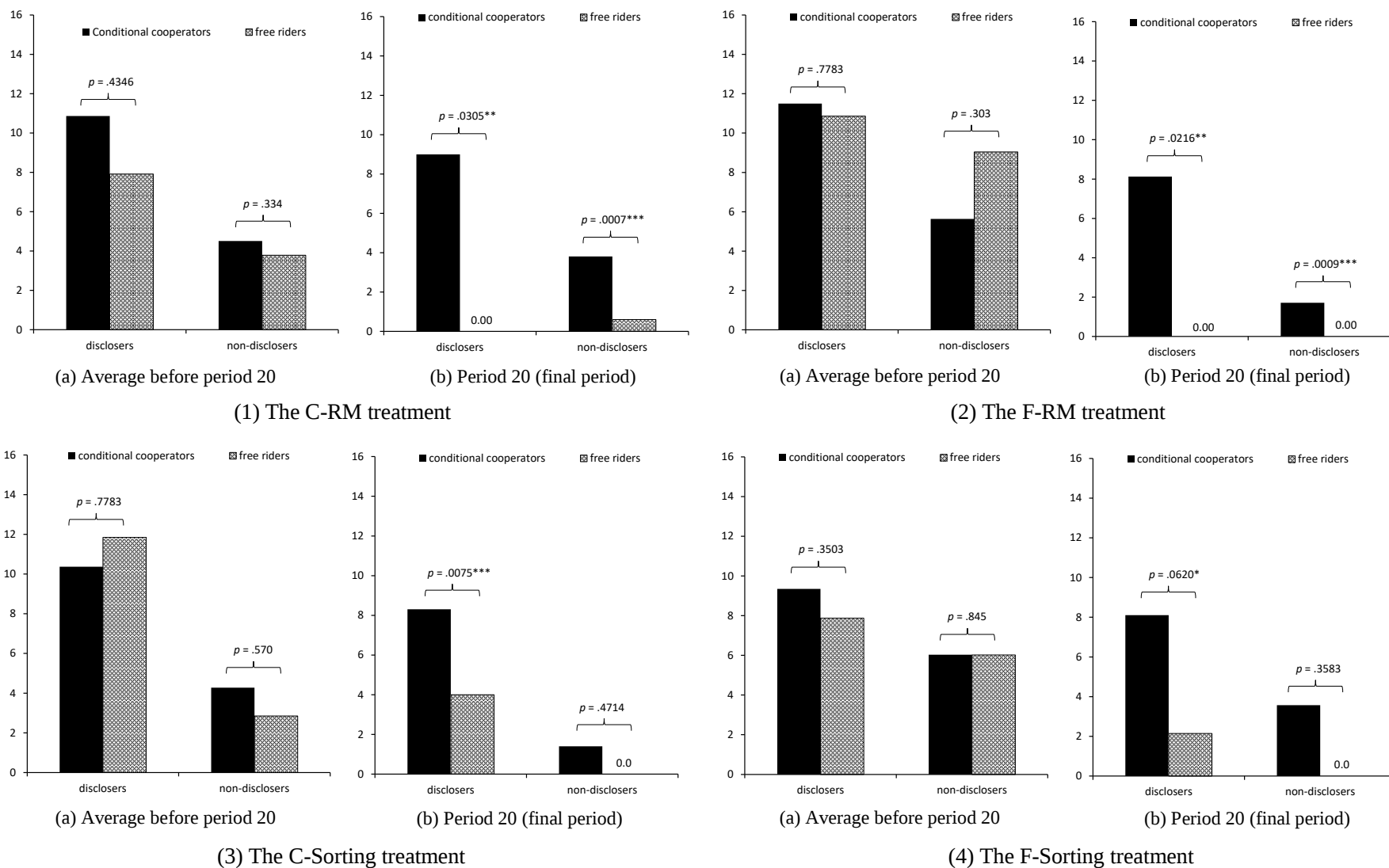
Result 8: (i) *Free riders contributed very little in period 20, unlike conditional cooperators.* (ii) *However, the free riders contributed almost the same amounts as the conditional cooperators in periods before period 20.*

One may wonder why some conditional cooperators did not disclose their states. One possibility is that they were pessimistic about the peers' behaviors. As shown in Fig. C.5, conditional cooperators who less frequently disclosed their states formed lower beliefs on their matched peers' contribution behaviors, compared with the same types who frequently disclosed.³⁶ However, this is suggestive evidence only because the subjects' average beliefs do not differ significantly by the frequency of disclosure (also see Appendix Table C.4).

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 disclose their states contributed larger amounts when matched with disclosers rather than non-disclosers, but they still contributed large amounts toward non-disclosers (Appendix Fig. C.4). 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 the states (each panel (b) of Fig. C.4). If we accept that free riders are strategically-minded, this behavior by masked free riders may have been driven by an attempt not to discourage disclosers' willingness to contribute for future exploitative purposes.

³⁶ I acknowledge that the direction of causality may be the opposite: subjects' beliefs on the disclosers' strong cooperation behaviors may have been formed by their experiences.

Fig. 5: Average Contribution Amounts by Disclosure Decision



Notes: Each panel (a) reports average contribution amounts across all periods but periods 1 and 20 by disclosure decision. See Appendix Table C.3 for the calculations of 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.

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

Consistent with the theoretical discussions, disclosers' contribution amounts were higher in the C-Sorting than in the F-Sorting treatment (Fig. 1, Fig. 2). This subsection considers the difference between the two sorting treatments in details.

Appendix Fig. C.6 reports 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 frequently selected to disclose in the F-Sorting than in the C-Sorting treatment. Appendix Fig. C.7 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.³⁷ The percentages of subjects who disclosed more than or equal to ten times (except period 20), which I call the “frequent disclosers” hereafter, are 52.3% and 77.3% in the C-Sorting and F-Sorting treatments, respectively.³⁸ Fig. C.6 also indicates that a larger percentage of the frequent disclosers contributed small amounts in the F-Sorting than in the C-Sorting treatment.³⁹ These suggest that as discussed in Section 3, the presence of 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.⁴⁰

Result 9: *A significantly larger percentage of disclosers attempted to exploit high contributors in the F-Sorting than in the C-Sorting treatment.*

³⁷ 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 Fig. C.7). 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.2).

³⁸ The two percentages are significantly different (two-sample test of proportions, two-sided p -value = .0141).

³⁹ The percentages of the frequent disclosers that 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).

⁴⁰ 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. One may think 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 we in fact have an effect of having a positive disclosure cost; however, the presence of disclosed information is crucial for performance in the C-Sorting treatment. See Appendix D for the detail.

6. Discussions and Conclusions

This study demonstrated that given an option to disclose the past, a large number of individuals did disclose their past action choices and that voluntary information disclosure helps strengthen cooperation if disclosure is free. However, individuals were deterred from disclosing if disclosure involved a cost, even small, and if it did not ensure a matching with another discloser. This is surprising, considering the small size of the disclosure cost involved – only 1 ECU (2.5 pence).⁴¹

However, 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 interactions (e.g., Tadelis and Zettelmeyer, 2015). Nevertheless, we found that 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 private cost.

Aspect (a) may be caused by subjects' limited cognitive ability. For example, Li and Schipper (2018), in the context of a simpler 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 possibility also reinforces my conclusion that mandatory disclosure is superior in our setup, since it decreases subjects' cognitive loads. What players need to evaluate in the mandatory system is the peers' responses to the value of own states S only.

One area in which the government often focuses on enforcing mandatory disclosure is health and safety (Dranove and Jin, 2010). Our results imply that implementing some form of

⁴¹ Kamei (2017) showed that people are less likely to hide own identity if doing so is costly than if it is costless in an infinitely repeated prisoner's dilemma game with random matching. The hiding cost accounted for 40% of the deviation gain of the stage game payoff matrix in the experiment. As discussed, the fact that the disclosure cost in the present study was very small suggests that there may be a discontinuity in people's disclosure behaviors between zero and positive costs.

mandatory disclosure by private sectors (e.g., mandatory disclosure of all reviews and comments on the online platforms) would be complementary to help enhance economic efficiency if people's opportunistic behaviors are liable to undermine the community norms. There may be, however, unwanted side effects, such as gaming behaviors among users (e.g., Dellarocas, 2003).

This paper is also related to a debate on the role of disclosed information in the empirical literature. While some research has documented that disclosed information has no impact, for example, in price advertising on price dispersion (Milyo and Waldfogel, 1999) and on the quality of child-care services (Chipty and Witte, 1998), other research has shown its positive impact, for example, of restaurant hygiene grade on the improvement of health outcomes (Jin and Leslie, 2003) and of photographs on traded prices in the eBay Motors (Lewis, 2011). Jin and Leslie (2003) argue that the impact of disclosed information may be large if the users would otherwise not be able to obtain that information easily. They argue in the context of hygiene that: “mandatory disclosure of hygiene grade leaves restaurants with no choice about whether to display the information. ... the grade cards may eliminate a substantial search cost for customers.” In my setup, players had no substitutes for the disclosed information except for their own experiences, whose aspect may lead to the positive impact of disclosed information.

Finally, I note that our findings also have implications for the design of information systems in other markets, such as the private rental housing market. The relevant authority sometimes requires disclosure of agents' past behaviors. For example, the mayor of London (Mr. Sadiq Khan) recently introduced the Rogue Landlord and Agent Checker in the Greater London area. All boroughs and the City of London have participated in the Checker since May 2018, which provides publicly available information about private landlords and real estate agents that have previously been prosecuted or fined. The introduction of this system may prevent certain landlords from engaging in anti-social behaviors, as evidenced in my study. There is no longer an option for landlords to conceal their anti-social behaviors and interactions between landlords and tenants are not one-shot.

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.
- Chipty, T. and Witte, A.D. "Effects of Information Provision in a Vertically Differentiated Market." NBER Working Paper 6493, 1998.
- Dellarocas, C. "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms." *Management Science*, Vol. 49 (2003), pp. 1407-1424.
- 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.
- Dufwenberg, M. and Kirchsteiger, G. "A theory of sequential reciprocity." *Games and Economic Behavior*, Vol. 47 (2004), pp. 268-298.
- Ehrhart, K.-M. and Keser, C. "Mobility and Cooperation: on the Run." CIRANO working paper 99s-24, 1999.
- Embrey, M., Fréchette, G. and Yuksel, S., "Cooperation in the Finitely Repeated Prisoner's Dilemma." *Quarterly Journal of Economics*, Vol. 133 (2018), pp. 509-551
- 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.
- 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.
- 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.
- Jin, G. and Leslie, P. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." *Quarterly Journal of Economics*, Vol. 118 (2003), pp. 409-451.
- 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.
- King, R. and Wallin, D. "Market-induced information disclosures: An experimental markets investigation." *Contemporary Accounting Research*, Vol. 8 (1991), pp. 170-197.
- 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.
- Lewis, G. "Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors." *American Economic Review*, Vol. 101 (2011), pp. 1535-1546.

- Li, Y. and Schipper, B. “Strategic Reasoning in Persuasion Games: An Experiment.” working paper, 2018.
- Milyo, J. and Waldfogel, J. “The Effect of Price Advertising on Prices: Evidence in the Wake of 44 Liquormart.” *American Economic Review*, Vol. 89 (1999), pp. 1081-1096.
- 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.
- 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.

Not for Publication

Supplementary Appendix for: Kamei, Kenju

“Voluntary Disclosure of Information and Cooperation in
Simultaneous-Move Economic Interactions”¹

Table of Contents

Appendix A: Some Experimental Procedure and Sample Instructions	p. 2
Appendix B: Theoretical calculations	p. 10
Appendix C: Additional Tables and Figures	p. 14
Appendix D: Results of the C-Sorting-N treatment	p. 29

¹ Correspondence: Kenju Kamei, Department of Economics and Finance, Durham University, Mill Hill Lane, Durham, DH1 3LB, United Kingdom. Email: kenju.kamei@gmail.com, kenju.kamei@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.

Please begin by filling in your conditional allocation schedule

If the other three members on average choose 0 points	
If the other three members on average choose 1 point	
If the other three members on average choose 2 points	
If the other three members on average choose 3 points	
If the other three members on average choose 4 points	
If the other three members on average choose 5 points	
If the other three members on average choose 6 points	
If the other three members on average choose 7 points	
If the other three members on average choose 8 points	
If the other three members on average choose 9 points	
If the other three members on average choose 10 points	
If the other three members on average choose 11 points	
If the other three members on average choose 12 points	
If the other three members on average choose 13 points	
If the other three members on average choose 14 points	
If the other three members on average choose 15 points	
If the other three members on average choose 16 points	
If the other three members on average choose 17 points	
If the other three members on average choose 18 points	
If the other three members on average choose 19 points	
If the other three members on average choose 20 points	

Continue

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.

Final: Your Unconditional Allocation Decision

Note that one out of the four group members will be randomly selected and the selected member's allocation points to the group account will be calculated based on his or her conditional allocation schedule. This unconditional allocation decision will be used if you are one of the three members whose conditional allocation schedule is not selected to be used.

Your endowment: 20

Your allocation to the group account:

- ☐ 0
- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7
- ☐ 8
- ☐ 9
- ☐ 10
- ☐ 11
- ☐ 12
- ☐ 13
- ☐ 14
- ☐ 15
- ☐ 16
- ☐ 17
- ☐ 18
- ☐ 19
- ☐ 20

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:

$$(\text{sum of points in your private account}) + 0.4 \times (\text{sum of points allocated by you and your group members to the group account})$$

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 ($= 1 \times .4$) 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, they moved on to Part 2. At the onset of Part 2, 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 do not 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 ($= 1 \times .8$) 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 ($= 14 \times 0.8$) points.

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

$$\text{(sum of points in your private account)} + 0.8 \times \text{(sum of points allocated by you and your counterpart to the joint account)}$$

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 $0.8 \times 1 = 0.8$ 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 $0.8 \times 2 = 1.6$ 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 $0.8 \times 1 = 0.8$ 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?

Appendix B: Theoretical Calculations

This part of the Appendix explains a material payoff maximizer i 's optimal control, assuming that no other material payoff maximizers mimic the behaviors of conditional cooperators. We can then derive i 's optimal control when some other free riders are mimickers by adjusting the probability that i is matched with those who act according to the conditional cooperation strategy (see the paper).

Notation:

p_{cc} : the percentage of conditional cooperators (those who disclose last-period contribution amounts and act on the conditional cooperation strategy)

p_{FR} : the percentage of free riders (those who contribute 0s unconditionally)

f : disclosure cost (= 1 in the C-RM and C-Sorting treatment; = 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: All free riders other than i behave as in Hypothesis 1 of the paper. Specifically, when disclosure is costly, no free riders other than i disclose the states. When disclosure is free, free riders other than i disclose the states with a probability of 50%.

As illustrated below, calculations indicate the following treatment differences:

- Provided that there is a sufficiently large percentage of conditional cooperators, the presence of a 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, in the C-RM and F-RM treatments $\tilde{p}_{cc} = .25$.
- Suppose that there is a sufficiently large numbers of conditional cooperators. A material payoff maximizer 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 disclosers are assured to be matched with another disclosers under the sorting condition.
- Subjects' cooperation behaviors in the Mandatory treatment are the same as those in the C-RM and F-RM treatments.

B.1. C-RM and F-RM treatments:

I 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 anything. This assumption means that

conditional cooperators contribute $S_1 = c_0$, while free riders contribute zero, in period 1. We can then set up the Hamiltonian function as in (#0):

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

This control problem is written as:

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

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

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

$$(\#1) \left\{ \begin{array}{l} \bullet \text{ For } t = 2, \dots, 20, \Delta \lambda_{i,t} = \lambda_{i,t+1} - \lambda_{i,t} = -\partial H_{t,i} / \partial S_{i,t} = -p_{cc} \cdot r \cdot 1_{\text{disclose},t} - \lambda_{i,t+1} \cdot (-1). \text{ In other words,} \\ \lambda_{i,t} = p_{cc} \cdot r \cdot 1_{\text{disclose},t}. \\ \bullet \text{ For } t = 21, \lambda_{i,21} = \partial \pi(i, 21) / \partial S_{i,20} = 0. \end{array} \right.$$

The optimal control for $\{c_{i,t}\}_{t \leq 20}$ 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:

$$(\#2) \left\{ \begin{array}{l} \bullet \text{ When } t < 20, \partial H_{t,i} / \partial c_{i,t} = -1 + r + \lambda_{i,t+1} = -1 + r + p_{cc} \cdot r \cdot 1_{\text{disclose},t+1}. \text{ This means } c_{i,t} = 20 \text{ if } r + p_{cc} \cdot r \cdot 1_{\text{disclose},t+1} > 1. \text{ Since } r = .8, \text{ this inequality reduces to: } p_{cc} \cdot 1_{\text{disclose},t+1} > .25. \text{ By contrast, } c_{i,t} = 0 \text{ if } -1 + r + p_{cc} \cdot r \cdot 1_{\text{disclose},t+1} < 0, \text{ or } p_{cc} \cdot 1_{\text{disclose},t+1} < .25. \\ \bullet \text{ When } t = 20, \partial H_{t,i} / \partial c_{i,t} = -1 + r < 0. \text{ Thus, } c_{i,20} = 0. \end{array} \right.$$

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

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

Because the disclosure cost is just one ECU, we have the same reputation building behaviors in the C-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

as we are assuming that $p_{cc} > .25$. Thus, if $p_{cc} > .25$, subject i always discloses and contributes 20 until period 19; and i discloses $S_{i,20}$ 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: *The presence of a positive cost does not affect the material payoff maximizer's incentive to mimic the behavior of conditional cooperator under random matching.*

(a) *If there are such a large fraction of conditional cooperators that $p_{cc} > .25$ in the population, i always 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 then selects $c_{i,20} = 0$.*

(b) *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 just randomly decides whether to disclose the state in each period in the F-RM treatment.*

B.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 percentage of subjects interacting with a conditional cooperator given that i discloses her state is larger, compared with the random matching environment. The probability is given by:

$$\text{Prob}[cc|disclose] = 1 \geq p_{cc}, \text{ when disclosure is costly.}$$

$$\text{Prob}[cc|disclose] = p_{cc}/[p_{cc} + 0.5p_{FR}] \geq p_{cc}, \text{ when disclosure is free.}$$

Here, notice that $\text{Prob}[cc|disclose]$ is larger in the C-Sorting than in the F-Sorting treatment.

Finding the optimal control path of a material payoff maximizer i in this case is identical to the case in Section B.1, except that the probability of being matched with a discloser in the Hamiltonian for $t > 1$ in Condition (#0) is $\text{Prob}[cc|disclose]$, instead of p_{cc} . By re-doing exactly the calculation process explained in Section B.1, we find the following reputation building behaviors:

Summary 2: *The presence of a positive cost increases the material payoff maximizer's incentive to mimic the behavior of conditional cooperator under sorting because the cost prevents other free riders from invading the set of disclosers. (a) If there is such a large fraction of conditional cooperators in the set of disclosers that $\text{Prob}[cc|disclose] > .25$, i always contributes 20 and then discloses the state 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 $\text{Prob}[cc|disclose] \geq p_{cc}$, when $\text{Prob}[cc|disclose] > .25$, the disclosure rate and average contribution amount are both higher in the C-Sorting (F-Sorting) than in the C-RM (F-RM) treatment. (b) By contrast, if $\text{Prob}[cc|disclose] < .25$ in the C-Sorting treatment, i always contributes nothing and hides her*

state. If $\text{Prob}[cc|disclose] < .25$ in the F-Sorting treatment, i always contributes nothing while i randomly decides whether to disclose.

B.3. Mandatory treatment:

Since we are assuming that no material payoff maximizers other than i mimic the behaviors of conditional cooperator, the analysis for the Mandatory treatment is the same as the one studied in Section B.1. In the Mandatory treatment, we have two types as below:

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

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

As the likelihood of being matched with a conditional cooperator is p_{cc} , this setup 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.*

Appendix C: Additional Tables and Figures

Table C.1. *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 among the treatments (see Fig. 2 for the average contribution amounts 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

[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

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

Table C.2. 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 $t \in \{2, 3, \dots, 19\}$; and 0 otherwise.

Treatment:	(a) C-Sorting			(b) C-RM		
	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 $\{= 2, 3, \dots, 19\}$	-.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 $\times i$'s period $t - 1$ matched partners' contribution amounts	---	.0023 (.0033)	---	---	.0037 (.0033)	.0010 (.0019)
Not-Disclosed-Last-Period Dummy $\times i$'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 or 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 or 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 defection observed in the experiment. Results are similar even if data from period 20 is 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 $p = .0062^{***}$

(Continued)

Independent variable:	Treatment:	(c) F-Sorting			(d) F-RM	
	All data	<i>i</i> disclosed in period $t - 1$	<i>i</i> did not disclose in period $t - 1$	All data	<i>i</i> disclosed in period $t - 1$	<i>i</i> did not disclose in period $t - 1$
	(c1)	(c2)	(c3)	(d1)	(d2)	(d3)
Period Number $\{= 2, 3, \dots, 19\}$	.00059 (.0038)	.0016 (.0025)	.012** (.0027)	-.0017 (.0023)	.0062** (.0015)	-.0017 (.0033)
The contribution amount of subject <i>i</i> in period $t - 1$	---	.022** (.0065)	.038* (.012)	---	.039*** (.0029)	.040** (.0074)
Disclosed-Last-Period Dummy $\times i$'s period $t - 1$ matched partners' contribution amounts	---	.0049** (.0014)	---	---	-.0024 (.0014)	-.0093 (.0041)
Not-Disclosed-Last-Period Dummy $\times i$'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 $p = .0027***$

Table C.3. *Subjects' Decisions to Contribute by Cooperation Type (supplementing Fig. 5 of the paper)*

Dependent variable: subject i 's contribution to the joint account in period t .				
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 (.95)	1.42 (2.50)	-4.41 (4.29)	-.59 (3.02)
(b) Conditional cooperator dummy $\times$ Disclosure dummy {= 1 if subject i disclosed his or her last-period contribution in period t ; and 0 otherwise}	3.32*** (1.03)	2.76*** (.50)	2.28* (1.22)	-.030 (1.81)
(c) Free rider dummy $\times$ Disclosure dummy	1.34 (2.79)	5.27*** (1.83)	-2.27 (2.97)	-2.73 (2.33)
(d) Conditional cooperator dummy $\times$ Period 20 dummy {= 1 for period 20; 0 otherwise}	-1.05 (1.19)	-3.44*** (1.06)	-3.88*** (1.12)	-2.59* (1.35)
(e) Free rider dummy $\times$ Period 20 dummy	-4.27*** (.45)	-3.78*** (1.02)	-13.9*** (2.59)	-5.3*** (.15)
(f) Variable (b) $\times$ Period 20 dummy	.11 (1.72)	1.35 (2.57)	-.64 (3.70)	1.52 (2.92)
(g) Variable (c) $\times$ Period 20 dummy	1.02 (2.92)	-2.95 (2.02)	5.11*** (1.48)	-4.78 (4.23)
Constant	4.38*** (.34)	4.71** (2.36)	11.74*** (3.63)	8.98*** (2.47)
# of observations	646	703	570	646
- Two-sided p-values for Wald tests for the following comparisons regarding <u>disclosers</u>: H_0: (a) + (b) = (c) [the null that the contribution amounts of conditional cooperators are the same as those of free riders <u>before</u> 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 <u>in</u> period 20] .0305** .0075*** .0216** .0620* - Two-sided p-values for Wald tests for the following comparison regarding <u>non-disclosers</u>: H_0: (a) + (d) = (e) [the null that the contribution amounts of conditional cooperators are the same as those of free riders <u>in</u> 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 was 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 1: Contribution decisions before period 20 (the final 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 2: Contribution decisions in period 20 (the final 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 $H_0: (a) + (d) = (e)$.

Table C.4. *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* discloses her state (in columns (1) and (3)) or does not disclose her state (in columns (2) and (4)) up to period 19

Independent variable:	C-Sorting treatment		F-Sorting treatment	
	disclosed (1)	did not disclose (2)	disclosed (3)	did not disclose (4)
The total number of periods in which subject <i>i</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, because of the strong end-game defection seen in the experiment, as in other regression analyses (except the one with treatment effects).

^{#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).

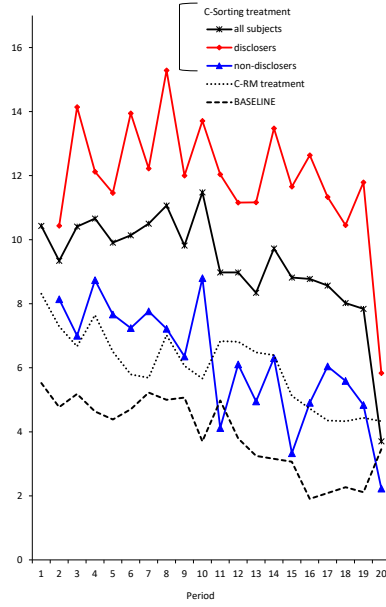
^{#3} One subject had never disclosed his or her state until period 19 in the F-Sorting treatment; thus, the total number of observations is 43 (= 44 – 1).

^{#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).

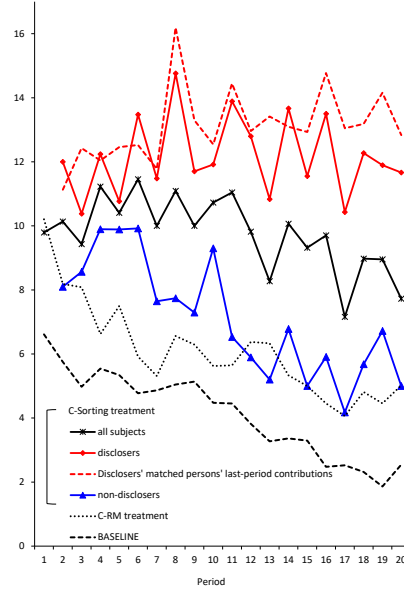
*, **, and *** indicate significance at the .10 level, at the .05 level and at the .01 level, respectively.

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

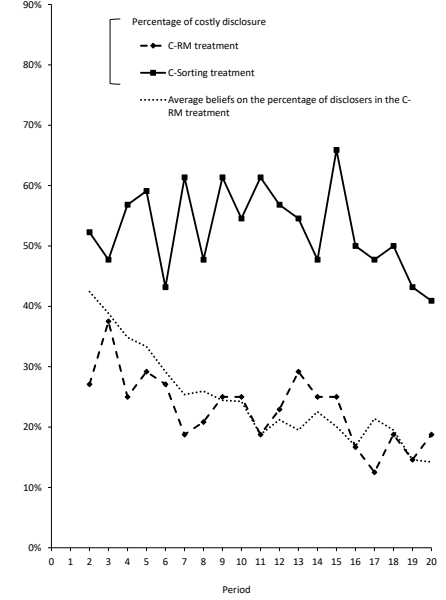
Fig. C.1. Period-by-Period Average Contributions and the Subjects' Disclosure Rates



(a) Avg. contribution

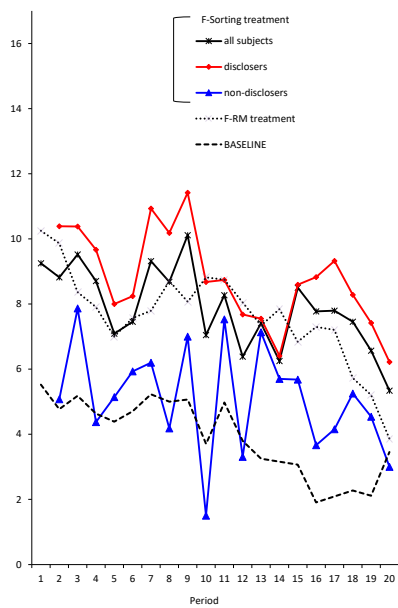


(b) Avg. belief on partners' contributions

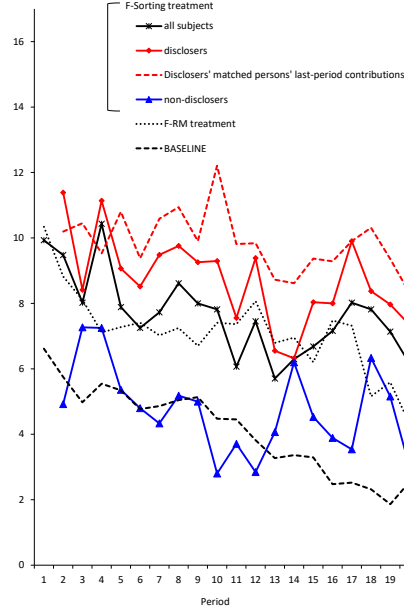


(c) Percentage of disclosers

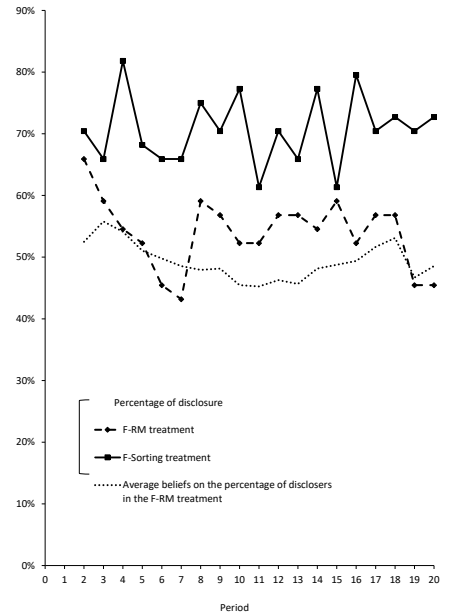
(I) C-RM and C-Sorting treatments



(a) Avg. contribution

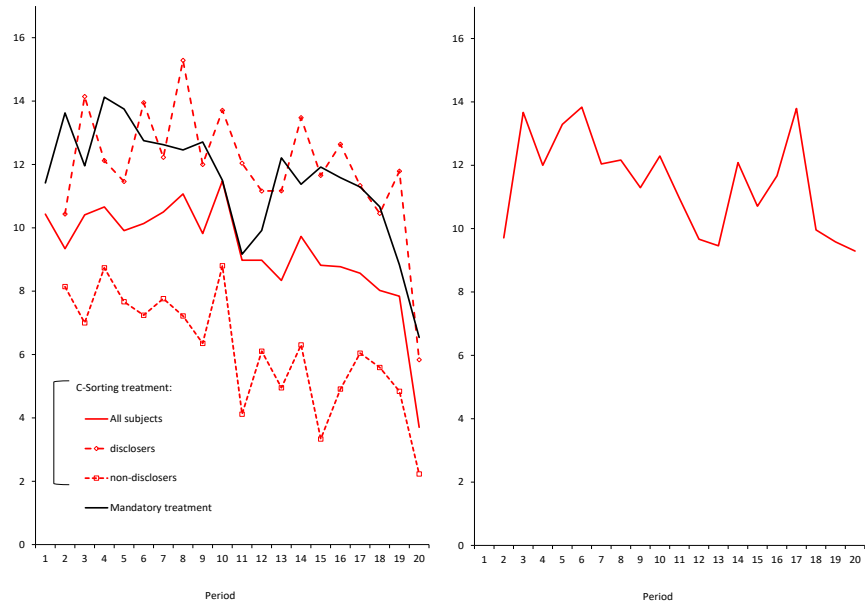


(b) Avg. belief on partners' contributions



(c) Percentage of disclosers

(II) F-RM and F-Sorting treatments



(a) Avg. contribution

(b) Avg. belief on partners' contributions

(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 treatment, respectively.

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

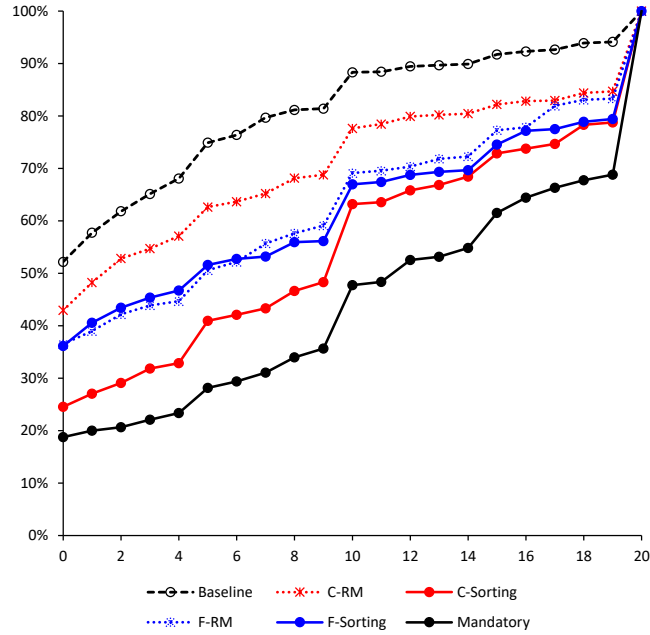
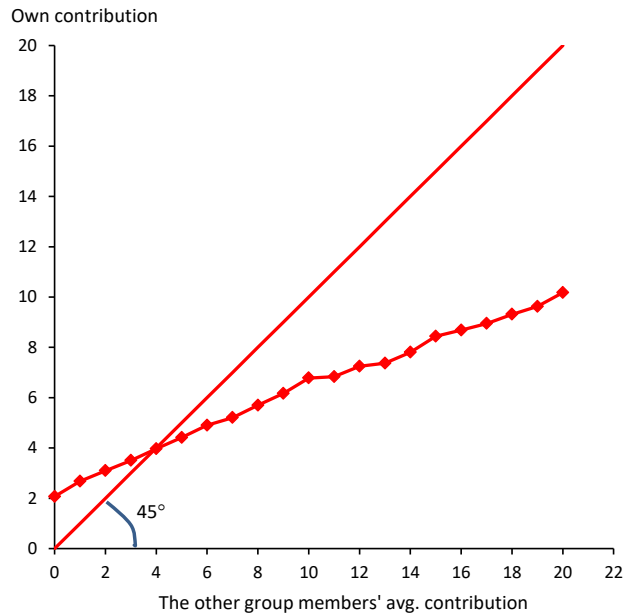


Fig. C.3. *Average Conditional Contribution Schedule*

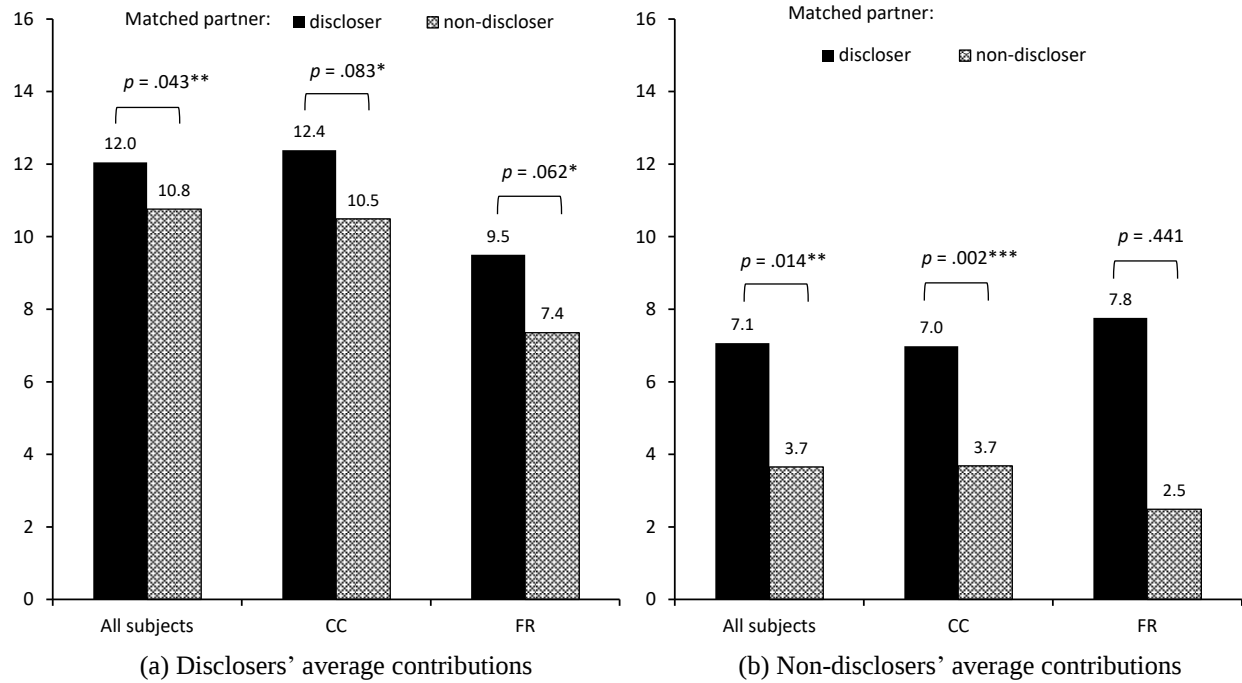


Notes: The above schedule is the average of all subjects' conditional contribution schedules.

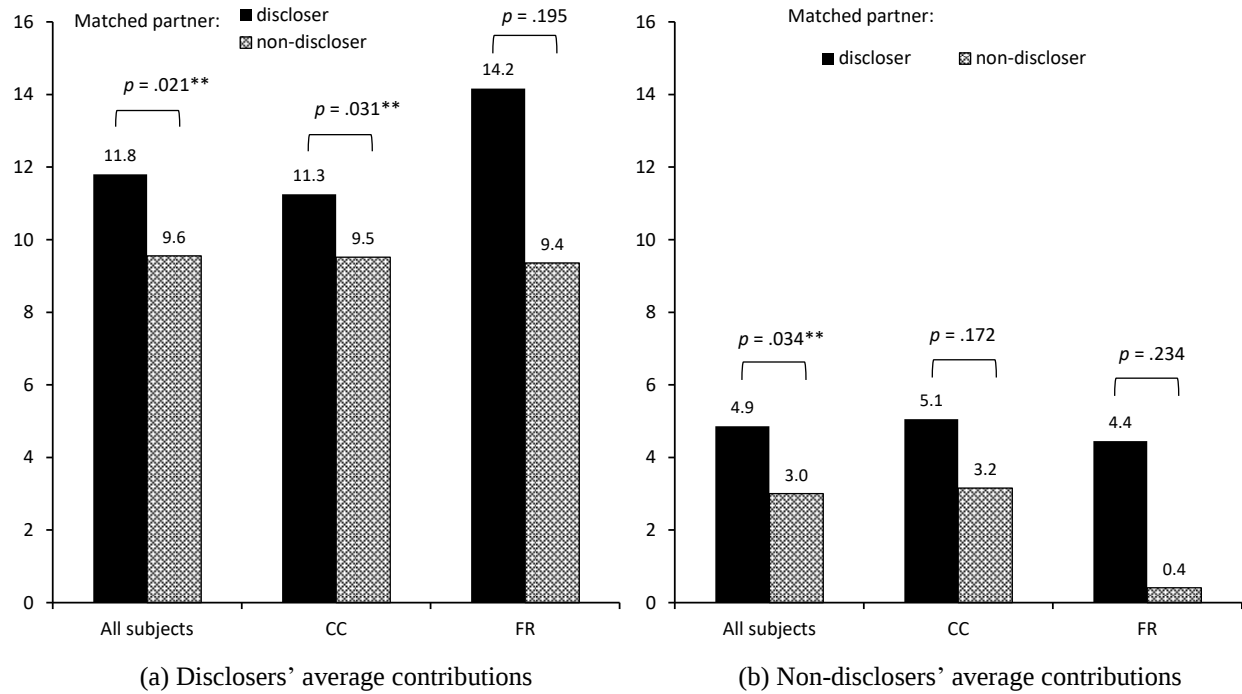
Remark: (1) The spearman's ρ correlation coefficient between own contribution amounts (y-axis) and the other group members' average contribution amounts (x-axis) is 1.0000 (p -value $< .001$).

(2) The slope 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 (two-sided p -value $< .001$).

Fig. C.4. Average Contribution Amounts by Disclosure Decision and by Conditional Contribution Type (supplementing Fig. 5 of the paper)



(1) The C-RM treatment



(2) The F-RM treatment

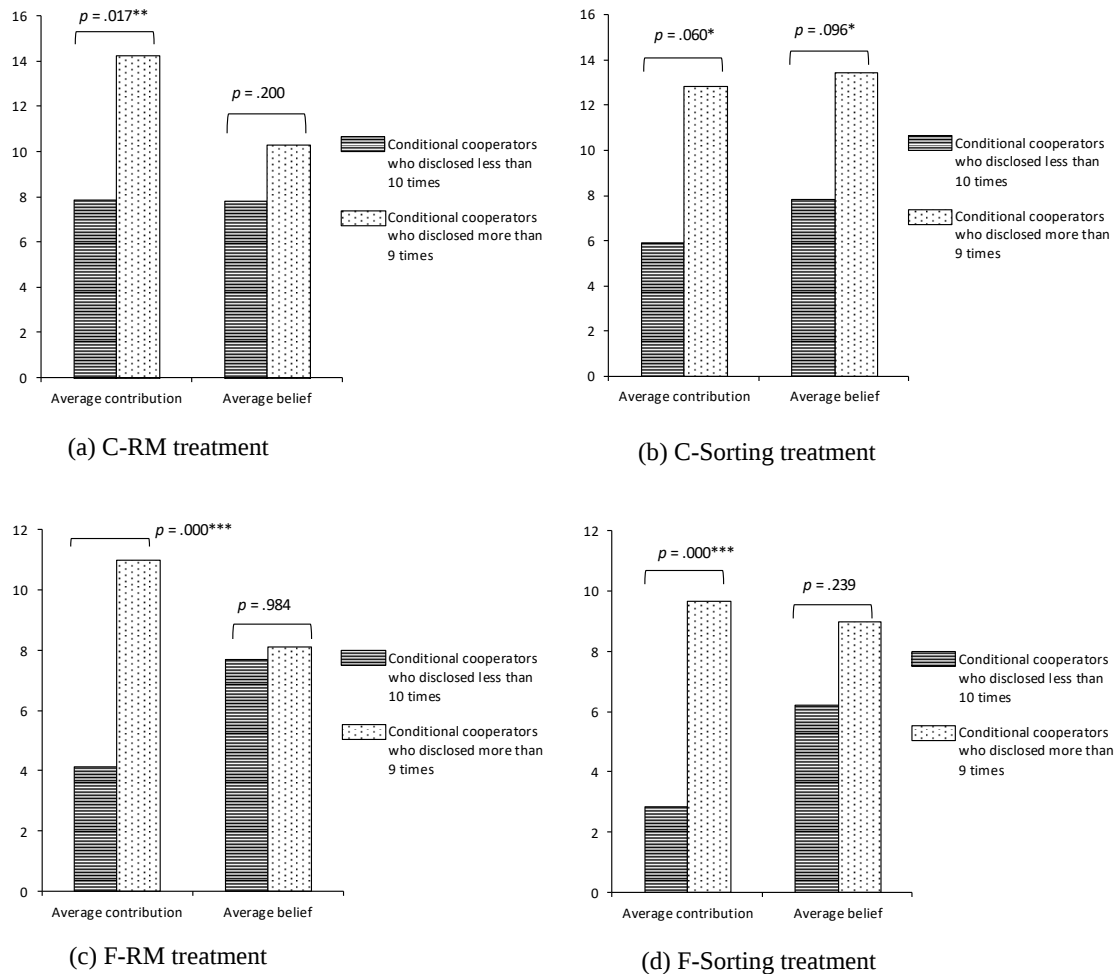
Notes: Average across all periods before period 20. p -values (two-sided) in the figure are based on individual fixed effect linear regressions with robust standard errors clustered by session.

Fig. C.5. Conditional Cooperators' Average Contributions and Beliefs when They Chose to Disclose, by the Frequency of Disclosure

We first classified conditional cooperators into those who chose to disclose more than nine times and those who chose to disclose less than ten times in the experiment. We then calculated the average contributions and beliefs for each category.

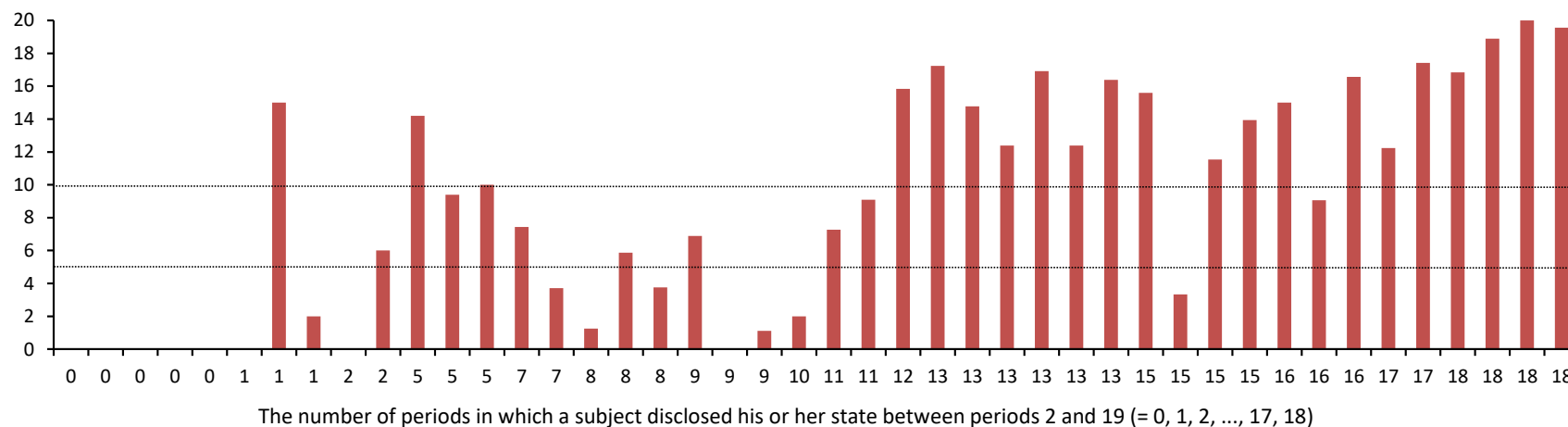
As shown below, conditional cooperators who less frequently disclosed formed lower beliefs on their matched peers' contribution behaviors, compared with the same types who frequently disclosed, although the differences in beliefs are not significant except for in the C-Sorting treatment.

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).

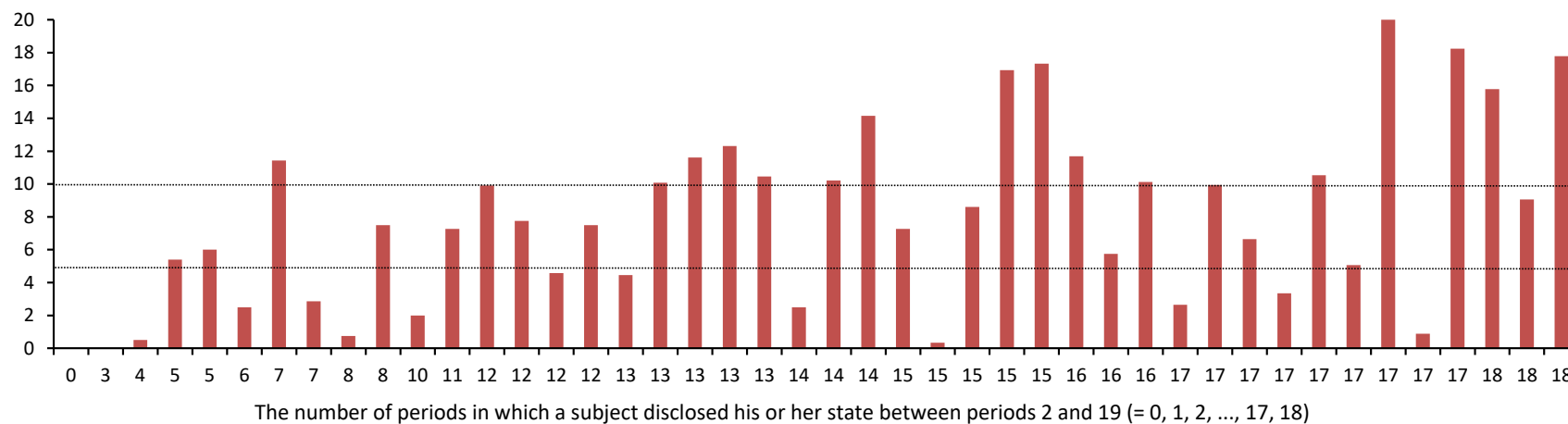


Notes: The figures show the average contributions and beliefs of conditional cooperators when they disclosed their states. The p -values in the figure are two-sided p -values, based on linear regressions (with robust standard errors clustered by session), where independent variables include a dummy which 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.

Fig. C.6. *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.7. Subject-by-Subject Patterns of Disclosure Decisions in the C-Sorting and F-Sorting treatments (supplementing Fig. C.6 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	1	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). Cases of non-disclosure are highlighted in yellow.

Remark: These two panels reveal different 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.

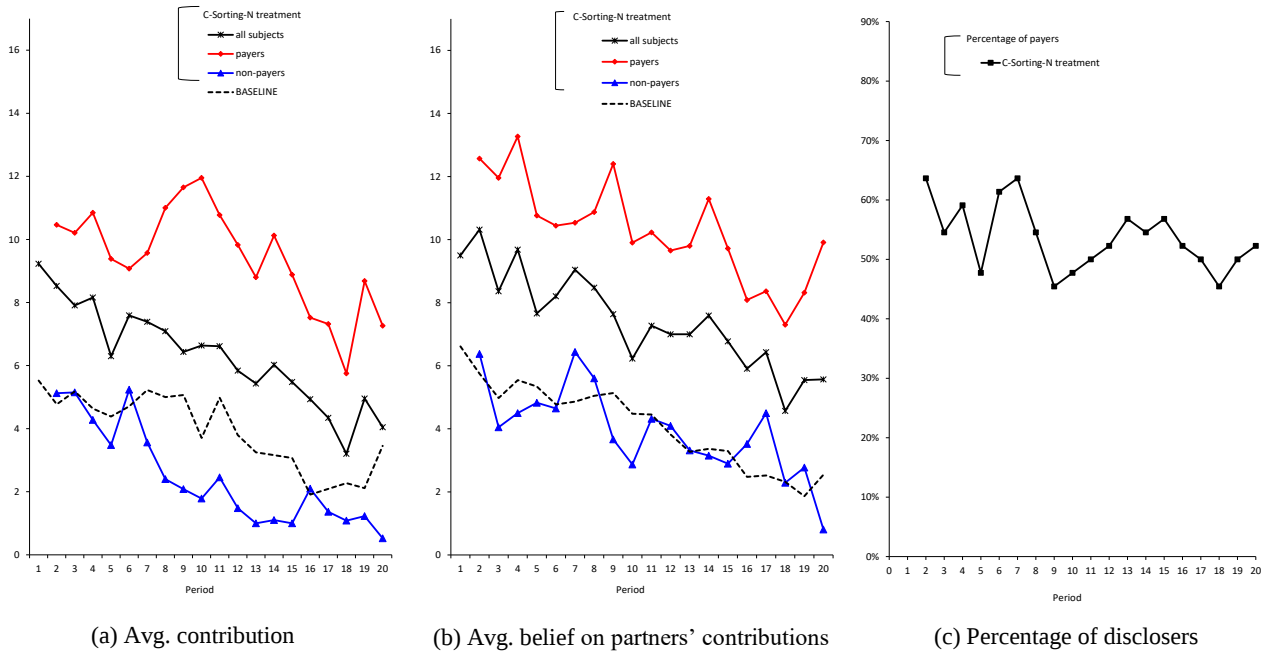
This observation can be partially confirmed by a regression analysis where the dependent variable is subject i 's decision to disclose (=1 if disclose; 0 otherwise) and the independent variables are (a) a dummy which equals 1(0) if i disclosed (did not disclose) in period $t - 1$, (b) the C-Sorting dummy 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. Using a random effect linear regression with robust standard errors clustered by session, I find that 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: Results of the C-Sorting-N treatment

The C-Sorting treatment achieved the highest efficiency, driven by the disclosers' strong cooperation behaviors, among the four main treatments. This was consistent with the theoretical analysis discussed in Section 3. This part of the Appendix is devoted to a further analysis for this observation.

The performance in 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, I conducted the "C-Sorting-N" treatment to study the relative importance of these two effects. In this treatment, subjects who paid one ECU was 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. D.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. D.2).

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



Nevertheless, the average contribution is 2.51 ECUs lower for the payers than for the disclosers in the C-Sorting treatment (Fig. D.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 D.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 (columns (1) and (2) in Table D.1). Qualitatively the same difference is found when average payoffs are compared between the C-Sorting-N and C-Sorting treatments (columns (3) and (4) in Table D.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 D.1: *The average contribution and average payoff were both significantly lower in the C-Sorting-N than in the C-Sorting treatment.*

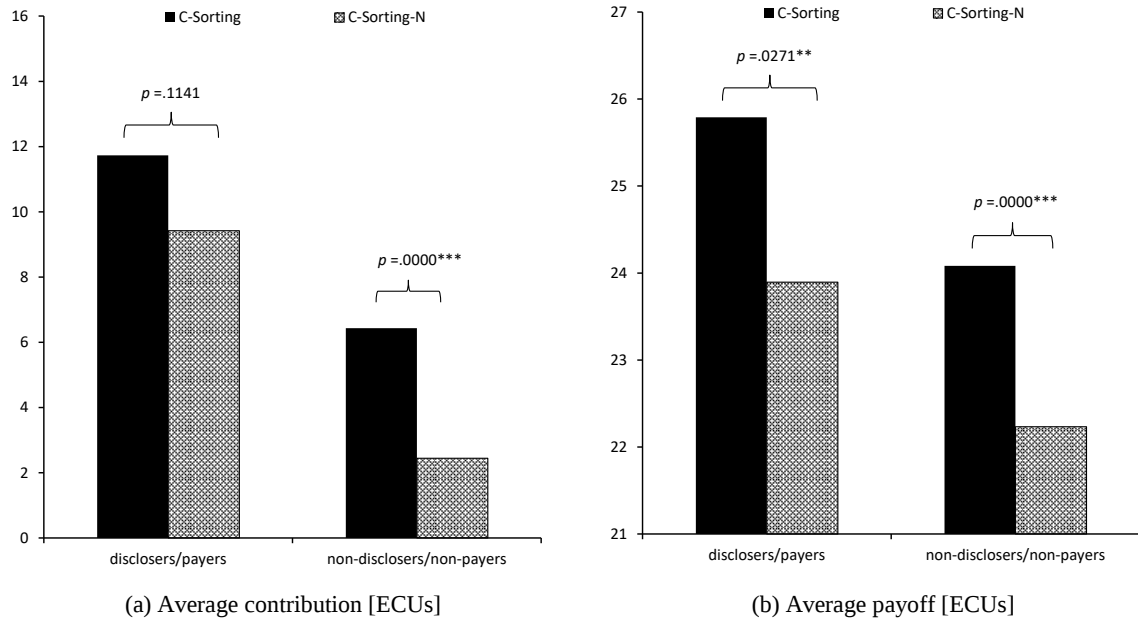
Table D.1. *The Impact of Paying a Cost without Information Disclosure in the C-Sorting-N treatment*

Dependent variable: Independent variable:	Average contributions in period t		Average payoffs in period t	
	(1)	(2)	(3)	(4)
C-Sorting dummy (#1)	5.19*** (1.40)	5.57*** (1.55)	2.60*** (.82)	2.79*** (.87)
C-Sorting-N dummy (#2)	2.21 (1.58)	3.01** (1.25)	.80 (.93)	1.21 (.75)
Period Number {= 2, 3, ..., 19, 20}	---	-.16*** (.039)	---	-.097*** (.024)
Constant	3.91*** (1.26)	5.70*** (1.08)	22.3*** (.76)	23.4*** (.65)
# of observations	228	228	228	228
p -value (two-sided) for Wald tests to the following hypotheses:				
$H_0: (\#1) = (\#2)$	.0082***	.0446**	.0038***	.0237**

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 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-N dummy equals 1 for the C-Sorting-N treatment; and 0 otherwise. In addition to the Period Number variable, the interaction term between each treatment dummy and the Period Number variable was added as a control in columns (2) and (4); we did not include them in this table to conserve space.

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

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



Notes: I first calculated session-average contributions and payoffs by disclosure decision. I then averaged them 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.