

# Indirect reciprocity in three types of social dilemmas

Mitsuhiro Nakamura<sup>1,\*</sup> and Hisashi Ohtsuki<sup>1</sup>

<sup>1</sup>*Department of Evolutionary Studies of Biosystems,  
The Graduate University for Advanced Studies (SOKENDAI),  
Hayama, Kanagawa 240-0193, Japan*

(Dated: March 19, 2014)

## Abstract

Indirect reciprocity is a key mechanism for the evolution of human cooperation. Previous studies explored indirect reciprocity in the so-called donation game, a special class of Prisoner's Dilemma (PD) with unilateral decision making. A more general class of social dilemma includes Snowdrift (SG), Stag Hunt (SH), and PD games, where two players perform actions simultaneously. In these simultaneous-move games, moral assessments need to be more complex; for example, how should we evaluate defection against an ill-reputed, but now cooperative, player? We examined indirect reciprocity in the three social dilemmas and identified twelve successful social norms for moral assessments. These successful norms have different principles in different dilemmas for suppressing cheaters. To suppress defectors, any defection against good players is prohibited in SG and PD, whereas defection against good players may be allowed in SH. To suppress unconditional cooperators, who help anyone and thereby indirectly contribute to jeopardizing indirect reciprocity, we found two mechanisms: indiscrimination between actions towards bad players (feasible in SG and PD) or punishment for cooperation with bad players (effective in any social dilemma). Moreover, we discovered that social norms that unfairly favour reciprocators enhance robustness of cooperation in SH, whereby reciprocators never lose their good reputation.

7 Keywords: evolutionary game theory; indirect reciprocity; Prisoner's Dilemma game; Snowdrift game; Stag  
8 Hunt game

---

\* Author for correspondence: [nakamuramh@soken.ac.jp](mailto:nakamuramh@soken.ac.jp)

## 9 I. INTRODUCTION

10 In everyday life, your social image influences what you obtain. Helping someone raises  
11 your reputation in your community and others help you later when required. This is called  
12 indirect reciprocity, a key mechanism for explaining the evolution of cooperative behav-  
13 ior among unrelated individuals [1–3]. Indirect reciprocity based on reputation has been  
14 extensively investigated for decades through numerous theoretical studies [4–23] and exper-  
15 imental tests [24–31]. The global success of humans in the past was partially dependent on  
16 the establishment of indirect reciprocity, as it was used to explore for more suitable part-  
17 ners for effective economic exchange instead of maintaining closed transactions in inefficient  
18 relationships [4, 32].

19 One important feature of indirect reciprocity is that it endogenously provides an incentive  
20 for actors to reward or punish other community members, which is achieved by controlling  
21 the actors’ reputations that lead to the future rewards or punishments for the actors them-  
22 selves. We can imagine numerous possibilities of rules to control the reputation of actors  
23 that behave differently in various social contexts; such rules are called social norms [4, 10].  
24 Some promising norms can stabilize cooperation in indirect reciprocity, but others cannot.  
25 Previous studies have systematically obtained successful social norms in Prisoner’s Dilemma  
26 scenarios when the reputation information is well-shared in a population [10, 11], when it  
27 belongs to each individual [9], with the presence of costly punishment [13], with incom-  
28 plete reputation information [19], with multiple reputation states [22], and with group-level  
29 reputations [21].

30 Most of the previous studies have investigated social norms for the so-called donation  
31 game, a variant of Prisoner’s Dilemma with unilateral decision making [33]. In the dona-  
32 tion game, two individuals called donor and recipient participate in and only the donor  
33 can decide whether or not to help the recipient, *i.e.*, whether to benefit the recipient by  
34 making an investment. Because the donation game focuses on the unilateral behavior of a  
35 donor, it ignores many aspects that exist in reality. One such aspect is that the donation  
36 game is merely an instance of various social dilemmas. Reputation systems would also play  
37 an important role in various simultaneous-move games such as Snowdrift, Stag Hunt, and  
38 general Prisoner’s Dilemma games. In these games, social norms may depend not only on  
39 an actor’s choice but also on his/her co-player’s choice. For example, how should we define

40 goodness when an actor defects against a bad co-player that unexpectedly cooperates with  
41 the actor? Should the actor's defection be justified, even if the co-player shows reformation?  
42 Moreover, individuals could infer that a focal player's reputation should be bad when the  
43 player received punishment from another player who had established a high reputaion. Can  
44 such possibility be stable in evolutionry scenarios? To the best of our knowledge, although  
45 two previous studies have investigated games other than the donation game, they have not  
46 done so exhaustively and not clarified the general characteristics of social norms for the  
47 simultaneous-move games [4, 18].

48 The present study is directed towards completely exploring reputation systems in  
49 simultaneous-move games that comprise more extensive social situations than those in  
50 the donation game. We discover that diverse social norms stabilize reciprocation and realize  
51 cooperative and stable populations. These successful social norms vary for different types  
52 of social dilemmas. To suppress cheating in Prisoner's Dilemma and Snowdrift games, these  
53 norms have a common characteristic such that defection against good players is regarded as  
54 bad irrespective of the co-player's action. However, in the Stag Hunt game, defection against  
55 good players may be allowed, whereas social norms that unfairly favour reciprocators are  
56 required to achieve robustness of reciprocation; under these norms, reciprocators never lose  
57 their good reputation. It is also imperative to punish unconditional cooperators that help  
58 anyone, because they blindly support cheaters [7, 8]. There are two mechanisms to restrain  
59 unconditional cooperation. One method is to avoid distinguishing between cooperation and  
60 defection towards bad players, in which case unconditional cooperators pay an extra cost of  
61 helping bad players while reciprocators do not. The other method is to regard cooperation  
62 with a bad player as a bad deed, in which case unconditional cooperators are explicitly  
63 punished. We discover that the former mechanism is feasible in Prisoner's Dilemma and  
64 Snowdrift games, whereas the latter works for all three social dilemmas.

## 65 II. MODEL

66 We consider a large, well-mixed population in which players from time to time play a  
67 symmetric two-player simultaneous-move game. In a one-shot game, two players are sampled  
68 from the population in a uniform random manner. Each player selects an action, which is  
69 either cooperation (C) or defection (D). There are four possible outcomes of the game for

70 a player: both players select C (the outcome is called reward; R), the focal player selects C  
71 and his/her co-player selects D (sucker; S), the focal player selects D and his/her co-player  
72 selects C (temptation; T), and both players select D (punishment; P). The payoff matrix of  
73 the game is given by

$$\begin{array}{cc} & \begin{array}{cc} \text{C} & \text{D} \end{array} \\ \begin{array}{c} \text{C} \\ \text{D} \end{array} & \begin{bmatrix} 1 & S \\ T & 0 \end{bmatrix}, \end{array} \quad (1)$$

74 where the payoff of the focal player is 1,  $S$ ,  $T$ , or 0 when the outcome is R, S, T, or P,  
75 respectively. Figure 1 illustrates the outcomes of competitions (*e.g.*, replicator dynamics)  
76 between cooperators and defectors for the three types of social dilemmas contained in the  
77 payoff matrix (1) [33–35]. In a two-dimensional payoff space, the region defined by  $T > 1 >$   
78  $S > 0$  yields a Snowdrift game (SG) that has one stable internal equilibrium at which the  
79 fraction  $S/(S + T - 1)$  of players are cooperators and the rest are defectors. The region  
80  $T > 1 > 0 > S$  yields a Prisoner’s Dilemma game (PD) that has a unique stable equilibrium  
81 at which defectors dominate the population. It should be noted that the donation game,  
82 where the sum of the payoffs of outcomes S (one-sidedly paying cost of helping) and T (one-  
83 sidedly enjoying benefit of being helped) is always equal to the payoff of outcome R (both  
84 paying cost and enjoying benefit), is projected onto a half-line  $S + T = 1$  ( $T > 1$ ) in the  
85 payoff space (solid red line in Fig. 1); the PD game defined here is more general than the  
86 donation game. The region  $1 > T > 0 > S$  yields a Stag Hunt game (SH) that has two pure  
87 stable equilibria at which cooperators and defectors each dominate the population. Because  
88 there is no dilemma when  $1 > T > 0$  and  $1 > S > 0$ , we do not study this trivial region.

89 We employ a binary reputation model in which reputation states are either good (G) or  
90 bad (B) (*e.g.*, Ref. [6]; see Refs. [33, 36, 37]). In a one-shot game, each of the two players  
91 selects an action (*i.e.*, C or D), which is a response to each co-player’s reputation (*i.e.*, G or  
92 B). A rule that specifies when to use which action is called an action rule, and it is denoted  
93 by  $a$ . There are four possible action rules. A reciprocator cooperates with a good co-player  
94 and defects against a bad co-player, *i.e.*,  $a(\text{G}) = \text{C}$  and  $a(\text{B}) = \text{D}$ . An unconditional  
95 cooperators always cooperates ( $a(\text{G}) = a(\text{B}) = \text{C}$ ) while an unconditional defector always  
96 defects ( $a(\text{G}) = a(\text{B}) = \text{D}$ ). A ‘contrary’ player cooperates with a bad co-player and defects  
97 against a good co-player ( $a(\text{G}) = \text{D}$  and  $a(\text{B}) = \text{C}$ ). Hereafter, we denote reciprocators,  
98 unconditional cooperators, unconditional defectors, and contrary players by CD, CC, DD,

99 and DC, respectively.

100 After a one-shot game, each participant of the game receives a new reputation that is  
101 determined by a social norm according to the outcome of the game (R, S, T, or P) and each  
102 co-player's reputation (G or B). Note that in our model, every member in a population has  
103 the same opinion about a player's reputation, which is attained through public information  
104 sharing [11, 36, 37]. Table I shows an example of a social norm under which a player receives  
105 a bad reputation only when he/she plays with a good co-player and the outcome is T or  
106 P, *i.e.*, whenever the player selects defection against a good co-player. This social norm  
107 was called simple standing in previous studies [12]. Because a social norm is specified by  
108 inserting G or B into the eight placeholders in a  $4 \times 2$  table, there are  $2^{4 \times 2} = 256$  possible  
109 norms.

110 We introduce errors in assessments with which a player is assigned an opposite reputation;  
111 if a player is assessed as good (bad), with a small probability  $\mu$ , the player receives a  
112 bad (good) reputation [9, 10]. The models of indirect reciprocity generally consider errors  
113 not only in observers' assessments but also in players' taking actions [38]. Nevertheless,  
114 because the difference between the two kinds of errors usually does not change the results  
115 qualitatively when assuming public information sharing (see, *e.g.*, Refs. [13, 20]), we only  
116 introduce errors in assessments.

### 117 III. METHODS

118 Our aim is to obtain desirable social norms that achieve cooperative and stable popula-  
119 tions of reciprocators in different social dilemmas. To do so, we verify whether each candidate  
120 of the 256 social norms satisfies the following criteria in each of three social dilemmas, SG,  
121 PD, and SH.

122 **Goodness:** The population of reciprocators develops mutual cooperation except for defec-  
123 tion caused by assessment errors.

124 **Stability:** The population of reciprocators is stable against any invasion by rare mutants  
125 (either CC, DD, or DC players).

126 Because the population of unconditional cooperators is stable in SH (see Fig. 1), one might  
127 wonder why we bother to need reciprocators and reputation systems to maintain cooperation.

128 One reason could be that reciprocators enhance robustness of cooperation. Therefore, for  
 129 SH, we additionally check the following criterion.

130 **Usefulness:** The population of reciprocators is more robust against an invasion by uncon-  
 131 ditional defectors than that of unconditional cooperators.

132 We extend the standard methods for indirect reciprocity in the donation game regime  
 133 (see, *e.g.*, Refs. [10, 19, 21]) to consider the simultaneous-move games, and introduce the  
 134 above three criteria. Table II summarizes the definitions of symbols used in this section.

### 135 A. Goodness

136 Consider a population in which all players adopt a unique action rule denoted by  $a$ . After  
 137 repeating the random matching games sufficiently many times, the population reaches an  
 138 equilibrium in which the fraction of players that have good reputations, denoted by  $p(\text{G})$ ,  
 139 satisfies

$$p(\text{G}) = \sum_{r_{\text{focal}} \in \{\text{G}, \text{B}\}} \sum_{r_{\text{co}} \in \{\text{G}, \text{B}\}} p(r_{\text{focal}}) p(r_{\text{co}}) \phi(g(a(r_{\text{co}}), a(r_{\text{focal}})), r_{\text{co}}). \quad (2)$$

140 The right-hand side of Eq. (2) averages the probability with which a player receives a good  
 141 reputation,  $\phi(g(a(r_{\text{co}}), a(r_{\text{focal}})), r_{\text{co}})$ , when the player and his/her co-player have reputations  
 142  $r_{\text{focal}}$  and  $r_{\text{co}}$ , respectively. In Eq. (2),  $g(a(r_{\text{co}}), a(r_{\text{focal}}))$  represents the outcome of the game  
 143 when the focal player and his/her co-player select actions  $a(r_{\text{co}})$  and  $a(r_{\text{focal}})$ , respectively.  
 144 Since the assessments involve errors that occur with probability  $\mu$ ,  $\phi(\cdot, \cdot)$  is either  $1 - \mu$  or  
 145  $\mu$ .  $p(\text{B}) (= 1 - p(\text{G}))$  represents the fraction of bad players. By solving Eq. (2), we obtain

$$p(\text{G}) = \begin{cases} \frac{B - \sqrt{B^2 - 4AC}}{2A} & (A \neq 0) \\ \frac{C}{B} & (A = 0), \end{cases} \quad (3)$$

where

$$A = \phi(g(a(\text{G}), a(\text{G})), \text{G}) + \phi(g(a(\text{B}), a(\text{B})), \text{B}) \\ - \phi(g(a(\text{B}), a(\text{G})), \text{B}) - \phi(g(a(\text{G}), a(\text{B})), \text{G}), \quad (4a)$$

$$B = 1 + 2\phi(g(a(\text{B}), a(\text{B})), \text{B}) \\ - \phi(g(a(\text{B}), a(\text{G})), \text{B}) - \phi(g(a(\text{G}), a(\text{B})), \text{G}), \quad (4b)$$

and

$$C = \phi(g(a(B), a(B)), B). \quad (4c)$$

146 We are interested in whether a homogeneous population of reciprocators achieves coop-  
 147 eration. In a population of reciprocators (*i.e.*, CD players), the frequency of cooperation  
 148 clearly equals the fraction of good players. Therefore, under a social norm, we expand the  
 149 obtained  $p(G)$  by  $\mu$ , and when

$$p(G) = 1 - O(\mu), \quad (5)$$

150 we regard the social norm as satisfying the criterion of goodness.

## 151 B. Stability

152 The expected payoff of a resident player in a homogeneous population of players adopting  
 153 an action rule  $a$  is given by

$$f(a|a) = \sum \sum p(r_{\text{focal}})p(r_{\text{co}})\psi(g(a(r_{\text{co}}), a(r_{\text{focal}}))), \quad (6)$$

154 where  $\psi(g(a(r_{\text{co}}), a(r_{\text{focal}})))$  determines a player's payoff for each outcome of the game, *i.e.*,  
 155  $g(a(r_{\text{co}}), a(r_{\text{focal}}))$ . Hereafter, we omit the ranges of the summations over  $r_{\text{focal}}$  and  $r_{\text{co}}$ .

156 We next consider that an infinitesimal fraction of mutant players adopting another action  
 157 rule  $b$  ( $\neq a$ ) invade the population. The fraction of mutants that have good reputations  
 158 satisfies

$$q(G) = \sum \sum q(r_{\text{focal}})p(r_{\text{co}})\phi(g(b(r_{\text{co}}), a(r_{\text{focal}})), r_{\text{co}}). \quad (7)$$

159 Equation (7) yields

$$q(G) = \frac{\delta - (\delta - \gamma)p(G)}{1 + \delta - \beta - (\alpha + \delta - \beta - \gamma)p(G)}, \quad (8)$$

where

$$\alpha = \phi(g(b(G), a(G)), G), \quad (9a)$$

$$\beta = \phi(g(b(B), a(G)), B), \quad (9b)$$

$$\gamma = \phi(g(b(G), a(B)), G), \quad (9c)$$

and

$$\delta = \phi(g(b(B), a(B)), B). \quad (9d)$$

160 The expected payoff of a mutant player is given by

$$f(b|a) = \sum \sum q(r_{\text{focal}})p(r_{\text{co}})\psi(g(b(r_{\text{co}}), a(r_{\text{focal}}))). \quad (10)$$

161 We define that the population of players adopting an action rule  $a$  is stable in a region  
162 of the payoff space (*i.e.*, SG, PD, or SH) if it satisfies

$$\Delta f \equiv f(b|a) - f(a|a) < 0 \quad \forall b \neq a \quad (11)$$

163 in all the area of the focused region.  $\Delta f$  is a function of  $\mu$  and thus, it can be expanded as

$$\Delta f = d_0 + \mu d_1 + O(\mu^2), \quad (12)$$

164 where  $d_k$  represents the series coefficient of  $k$ -th order when expanded by  $\mu$ . Because  $\Delta f$  is  
165 indeed at most of  $O(\mu)$  when the population satisfies the goodness criterion, we only need  
166 to check at most  $d_1$ . We consider that  $\Delta f < 0$  if

$$\begin{cases} d_0 < 0 & (d_0 \neq 0) \\ d_1 < 0 & (d_0 = 0 \text{ and } d_1 \neq 0) \end{cases} \quad (13)$$

167 holds true.

### 168 C. Usefulness

169 Because a homogeneous population of unconditional cooperators (*i.e.*, CC players) is  
170 stable in SH, even when a social norm satisfies the stability criterion, it is worse to adopt  
171 the CD action rule if the basin of attraction of CD players in competition with DD players  
172 is narrower than that of CC players. To examine this point, after detecting the social  
173 norms that satisfy the criteria of goodness and stability in SH, we numerically compare the  
174 basins of attraction of CC and CD players when they compete with DD players under those  
175 candidates. We select only the norms whereby CD players have larger basins of attraction  
176 than that of CC players in all the area of SH.

Here we consider a population that consists of players adopting either two action rules denoted by  $a$  and  $b$ . We denote by  $x$  the fraction of  $a$ -players; the fraction  $1 - x$  are  $b$ -players. We also denote by  $p(G)$  and  $q(G)$  the fractions of good players within  $a$ - and  $b$ -players, respectively. Note that the fraction of good players in the entire population equals



$xp(\text{G}) + (1-x)q(\text{G})$ .  $p(\text{G})$  and  $q(\text{G})$  are governed by the following time evolution:

$$\begin{aligned}\dot{p}(\text{G}) = & -p(\text{G}) + x \sum \sum p(r_{\text{focal}})p(r_{\text{co}})\phi(g(a(r_{\text{co}}), a(r_{\text{focal}})), r_{\text{co}}) \\ & + (1-x) \sum \sum p(r_{\text{focal}})q(r_{\text{co}})\phi(g(a(r_{\text{co}}), b(r_{\text{focal}})), r_{\text{co}}),\end{aligned}\quad (14a)$$

and

$$\begin{aligned}\dot{q}(\text{G}) = & -q(\text{G}) + x \sum \sum q(r_{\text{focal}})p(r_{\text{co}})\phi(g(b(r_{\text{co}}), a(r_{\text{focal}})), r_{\text{co}}) \\ & + (1-x) \sum \sum q(r_{\text{focal}})q(r_{\text{co}})\phi(g(b(r_{\text{co}}), b(r_{\text{focal}})), r_{\text{co}}).\end{aligned}\quad (14b)$$

We numerically solve  $\dot{p}(\text{G}) = \dot{q}(\text{G}) = 0$  in Eq. (14) and obtain the equilibrium values of  $p(\text{G})$  and  $q(\text{G})$  that satisfy  $\text{Tr } \mathcal{J} < 0$  and  $\det \mathcal{J} > 0$ , where  $\mathcal{J}$  is the Jacobian matrix of Eq. (14). Using them, the expected payoffs of  $a$ - and  $b$ -players, which depend on  $x$ , are given by

$$\begin{aligned}f(a|x) = & x \sum \sum p(r_{\text{focal}})p(r_{\text{co}})\psi(g(a(r_{\text{co}}), a(r_{\text{focal}}))) \\ & + (1-x) \sum \sum p(r_{\text{focal}})q(r_{\text{co}})\psi(g(a(r_{\text{co}}), b(r_{\text{focal}}))),\end{aligned}\quad (15a)$$

and

$$\begin{aligned}f(b|x) = & x \sum \sum q(r_{\text{focal}})p(r_{\text{co}})\psi(g(b(r_{\text{co}}), a(r_{\text{focal}}))) \\ & + (1-x) \sum \sum q(r_{\text{focal}})q(r_{\text{co}})\psi(g(b(r_{\text{co}}), b(r_{\text{focal}}))),\end{aligned}\quad (15b)$$

177 respectively.

178 When the competition between  $a$ - and  $b$ -players is bistable, the basin of attraction of  
179  $a$ -players is given by  $1 - x_a^*$ , where  $x_a^*$  is the critical fraction of  $a$ -players at which  $f(a|x_a^*) =$   
180  $f(b|x_a^*)$  holds true. Because the competitions between CC or CD players and DD players  
181 are indeed bistable in SH, to compare the basins, we only need to compare  $x_{\text{CC}}^*$  and  $x_{\text{CD}}^*$   
182 under each social norm. In case of the competition between CC and DD players, we easily  
183 obtain  $x_{\text{CC}}^* = S/(S + T - 1)$ . In case of the competition between CD and DD players, we  
184 fix  $a(\text{G}) = \text{C}$ ,  $a(\text{B}) = \text{D}$ , and  $b(\text{G}) = b(\text{B}) = \text{D}$  for Eqs. (14) and (15), set  $\mu = 0.01$  or  $0.1$ ,  
185 and identify  $x_{\text{CD}}^*$  using the bisection method [39]. For each social norm, we check whether  
186  $x_{\text{CD}}^* < x_{\text{CC}}^*$  holds true for all the payoff configurations  $(T, S) \in \{\epsilon, 2\epsilon, \dots, 1 - \epsilon\} \times \{-(1 -$   
187  $\epsilon), -(1 - 2\epsilon), \dots, -\epsilon\}$ , where we set  $\epsilon = 0.05$ .

188 After the above numerical examination, in Appendix A, we analytically verified whether  
189 the obtained norms (shown in Fig. 2(e)) satisfy the usefulness criterion.

## 190 IV. RESULTS

191 We found that, among the 256 candidates, twelve social norms shown in Fig. 2 satisfy the  
192 goodness, stability, and/or usefulness criteria for at least one of SG, PD, and SH. Among  
193 these twelve norms, eight satisfy the two criteria defined for SG and PD (Fig. 2(a, b1, b2))  
194 and six satisfy the three criteria defined for SH (Fig 2(b1, d1, d2)). The accurate conditions  
195 for the stability of reciprocators are listed in Appendix B.

### 196 A. Snowdrift and Prisoner’s Dilemma games

197 The four social norms shown in Fig. 2(a) satisfy the two criteria for SG and PD. A  
198 sufficient condition for the stability of reciprocators under these norms is given by  $T > 1$ .  
199 In these four norms, the assessments of an action towards a good co-player do not depend  
200 on the co-player’s action; cooperation with a good co-player is always regarded as good and  
201 defection against a good co-player is always regarded as bad. When a player encounters  
202 a bad co-player that selects cooperation (*i.e.*, outcome R or T), any action performed by  
203 the focal player is regarded as good. When a player encounters a bad co-player that selects  
204 defection (*i.e.*, outcome S or P), the assessment varies among the four norms.

205 The four social norms shown in Fig. 2(b1,2) also satisfy the two criteria for SG and  
206 PD. The condition for the stability of reciprocators under these four norms is given by  
207  $S < T$ . These four norms are different from those shown in Fig. 2(a) with regard to only the  
208 assessment such that mutual cooperation (*i.e.*, outcome R) with a bad co-player is regarded  
209 as bad.

210 Figure 2(c) extracts the common features of the eight norms in Fig. 2(a, b1, b2) that are  
211 successful in SG and PD. These norms claim that cooperation (*i.e.*, outcomes R and S) and  
212 defection (*i.e.*, outcomes T and P) towards good players should be regarded as good and  
213 bad, respectively, while one-sided defection (*i.e.*, outcome T) against bad players should be  
214 regarded as good.

### 215 B. Stag Hunt game

216 The four social norms shown in Fig. 2(d1,2) as well as those shown in Fig. 2(b1) satisfy the  
217 three criteria for SH. It should be noted that in SH, the four norms shown in Fig. 2(b1,2)

218 satisfy the goodness and stability criteria; however, only those in Fig. 2(b1) satisfy the  
 219 usefulness criterion. A sufficient condition for the stability of reciprocators under the four  
 220 norms in Fig. 2(d1,2) is given by  $S < T < 3/2$ , whereas the corresponding condition under  
 221 the two norms in Fig. 2(b1) is given by  $S < T$ . The four norms in Fig. 2(d1,2) are different  
 222 from those in Fig. 2(b1) with respect to the assessments such that either one-sided or mutual  
 223 defection (*i.e.*, outcome T or P) against a good co-player is regarded as good.

224 Figure 2(e) extracts the common features of the six norms shown in Fig. 2(b1, d1, d2) that  
 225 are successful in SH. These norms require that cooperation (*i.e.*, outcomes R and S) with  
 226 a good player and defection (*i.e.*, outcomes T and P) against a bad player are regarded as  
 227 good, *i.e.*, reciprocation should always be regarded as good. In addition, mutual cooperation  
 228 (*i.e.*, outcome R) with a bad player is regarded as bad.

## 229 V. INTUITIONS

230 From the twelve social norms obtained, we discovered that reputation systems are based  
 231 on different mechanisms to maintain indirect reciprocity. In this section, we provide ex-  
 232 planations for how a homogeneous population of reciprocators (*i.e.*, CD players) prevents  
 233 invasions by mutants that adopt the DD or DC action rules (Sec. VA) and the CC action  
 234 rule (Sec. VB).

### 235 A. Universality and an exception for excluding unconditional defectors and con- 236 trary players

237 Let us consider an invasion event in an error-free limit (*i.e.*,  $\mu \rightarrow 0$ ) under any successful  
 238 social norm. Here, most players (residents) adopt the CD action rule and an infinitesimal  
 239 fraction of players (mutants) adopt the DD or DC action rule. Because the social norm  
 240 satisfies the goodness criterion, most residents have good reputations, whereas we assume  
 241 that mutants have good and bad reputations with probabilities  $q(G)$  and  $q(B)$ , respectively.  
 242 In this population, a mutant is likely to play a game with a good CD resident. In the game,  
 243 the DD or DC mutant selects defection because the resident is of a good reputation, whereas  
 244 the CD resident selects cooperation or defection depending on the mutant's reputation.  
 245 Therefore, the outcome for the mutant is T (one-sided defection) when his/her reputation

246 is good, and P (mutual defection) when his/her reputation is bad. The expected payoff of  
 247 the mutant is  $q(\text{G}) \cdot T + q(\text{B}) \cdot 0 = q(\text{G})T$ , and that of the resident is clearly 1. The payoff  
 248 difference is thus  $\Delta f = q(\text{G})T - 1$ . The condition for stability against an invasion by the  
 249 mutants is  $\Delta f < 0$ , which is rewritten as

$$q(\text{G}) < \frac{1}{T}. \quad (16)$$

250 From Eq. (16), we see that there are two cases in which the mutants are suppressed. In one  
 251 case,  $q(\text{G})$  is sufficiently small, *i.e.*, the reputation of mutants is effectively damaged. This  
 252 policy is employed by the social norms in Fig. 2(c) that stabilize CD players in SG and PD.  
 253 They have a universal principle as per which, when a player plays with a good co-player,  
 254 cooperation (*i.e.*, outcome R or S) and defection (*i.e.*, outcome T or P) are regarded as good  
 255 and bad, respectively. Because of this principle, once a DD or DC mutant appears in the  
 256 population, he/she repeatedly encounters good players, selects defection, and receives bad  
 257 reputations. Intuitively, because the temptation of defection is considerably strong in SG  
 258 and PD (*i.e.*,  $T > 1$ ), defection against a good player should be accused.

259 In the other case,  $T$  is smaller than 1 and the inequality (16) is satisfied by any value of  
 260  $q(\text{G})$ . This is naturally met in SH and some of the social norms shown in Fig. 2(e) disregard  
 261 the reputation of defection against a good co-player (*i.e.*, outcome T or P). Intuitively  
 262 speaking, because defection against cooperation is simply irrational (*i.e.*,  $T < 1$ ) in SH,  
 263 there is no requirement to damage the reputations of unconditional defectors or contrary  
 264 players as punishment. However, to satisfy the usefulness criterion in SH, the reputation of  
 265 these mutants should be slightly damaged (see Appendix C); thus, the norms in Fig. 2(e)  
 266 have at least one pivot that assigns a bad reputation to defection, either one-sided or mutual  
 267 (*i.e.*, outcome T or P), against good players (see the ‘†’-ed pivots in Fig. 2(e)).

## 268 B. Diversity for excluding unconditional cooperators

269 In contrast, imagine that rare CC players invade a population of CD players. A CC mu-  
 270 tant is likely to encounter a good CD resident and always select cooperation. The expected  
 271 payoff of the mutant is  $q(\text{G}) \cdot 1 + q(\text{B}) \cdot S = q(\text{G}) + q(\text{B})S$ , and that of the resident is 1. The  
 272 payoff difference is thus  $\Delta f = [q(\text{G}) + q(\text{B})S] - 1 = -q(\text{B})(1 - S)$ . Because  $S < 1$  holds true  
 273 in all three social dilemmas, the condition for stability against an invasion by CC mutants,

274  $\Delta f < 0$ , is

$$q(\text{B}) > 0 \tag{17}$$

275 in the error-free limit (*i.e.*,  $\mu \rightarrow 0$ ). Equation (17) implies that a small yet non-erroneous re-  
276 duction of reputation suffices to suppress unconditional cooperators. However, by observing  
277 Fig. 2, it is evident that unconditional cooperators in most cases receive good reputations  
278 because selecting cooperation (*i.e.*, outcomes R and S) when one plays with a good co-player  
279 is always regarded as good under those norms. Since both CD and CC players generally have  
280 good reputations, the payoff difference between them is yielded by their different behaviors  
281 when they encounter rare bad players, who have erroneously received bad reputations.

282 In the four social norms shown in Fig. 2(a), when a CD or a CC player (both have good  
283 reputations) encounters a bad CD co-player, each selects defection and cooperation, while  
284 the CD co-player selects cooperation. As a result, both the focal CD and CC players receive  
285 good reputations. The payoff difference under these norms is, therefore, approximated by

$$\Delta f \propto 1 - T, \tag{18}$$

286 which yields the sufficient condition (shown before) for the stability of CD players against  
287 an invasion by CC mutants, *i.e.*,  $T > 1$ . During the above game sequence, the CD and  
288 CC players experience outcomes T and R, respectively, whereas their resultant reputations  
289 remain the same. Intuitively, if the temptation for defection is sufficiently large (*i.e.*, if  
290  $T > 1$ ) and an actor's behavior towards a bad player does not influence his/her reputation,  
291 defection is more rational than cooperation. This mechanism is feasible only in SG and PD.

292 In the eight social norms shown in Fig. 2(b1, b2, d1, d2), the adopted mechanism is  
293 different. When a focal (good) CD player encounters a bad CD co-player, he/she selects  
294 defection, whereas the co-player selects cooperation. As a result, the focal player maintains  
295 a good reputation. In the next round, the focal CD player encounters another good CD  
296 co-player, and mutual cooperation is achieved. The sum of payoffs in these two rounds is  
297  $T + 1$ . In contrast, when a focal (good) CC player encounters a bad CD co-player, both  
298 select cooperation, and the focal CC player receives a bad reputation. In the next round, the  
299 focal, bad CC player encounters a good CD co-player. The focal player selects cooperation,  
300 whereas the co-player selects defection. As a result, the focal CC player retrieves a good  
301 reputation. The sum of payoffs in these two rounds is  $1 + S$ . Thus, the payoff difference

302 between the CC and the CD players is approximated by

$$\Delta f \propto (1 + S) - (T + 1) = S - T, \quad (19)$$

303 which yields the condition (shown before) for the stability of CD players against an inva-  
304 sion by CC mutants, *i.e.*,  $S < T$ . Intuitively, these eight norms enforce defection against  
305 potentially harmful players, *i.e.*, bad players. Unconditional cooperators do not obey this  
306 enforcement and are punished for a moment. This mechanism is feasible in all three social  
307 dilemmas.

## 308 VI. DISCUSSION

### 309 A. Summary

310 We analyzed an extended model of indirect reciprocity in symmetric two-player simultaneous-  
311 move games that include three types of social dilemmas: Snowdrift (SG), Prisoner’s Dilemma  
312 (PD), and Stag Hunt (SH) games. We showed that twelve social norms achieve cooperative  
313 and stable populations of reciprocators that exclusively cooperate with good co-players  
314 (Fig. 2). These norms possess different characteristics for providing the stability to re-  
315 ciprocators in different payoff structures and in excluding mutants. In SG and PD, eight  
316 norms stabilize the populations of reciprocators (Fig. 2(c)). In SH, six norms stabilize the  
317 populations of reciprocators and also enable them to secure larger basins of attraction than  
318 unconditional cooperators in competition with unconditional defectors (Fig. 2(e)). Among  
319 them, only two norms are almighty such that they satisfy all the criteria in any type of social  
320 dilemmas (Fig. 2(b1)). These two norms are the variants of the so-called Kandori social  
321 norm, which is characterized as possessing enforcement of defection against bad players and  
322 is known to exhibit strong stability in previous models [4, 14, 16].

323 The twelve social norms implement mechanisms in diverse manners for detecting and  
324 punishing players that do not follow reciprocation. We confirmed a principle in SG and PD  
325 for preventing an invasion by defectors; cooperation (*i.e.*, outcomes R and S) and defection  
326 (*i.e.*, outcomes T and P) towards a good player should be regarded as good and bad,  
327 respectively. This principle is identical to one of the fundamental properties in the so-called  
328 ‘leading eight’ social norms that have been known to stabilize indirect reciprocity in the  
329 donation game regime [10, 11]. In the norms in Fig. 2(d1,2), either one-sided or mutual

330 defection (*i.e.*, outcome T or P) against a good player is regarded as good, and the defectors  
331 are not severely punished. This exception is only plausible in SH, because the temptation  
332 for defection is weak in SH (*i.e.*,  $T < 1$ ).

333 An invasion by unconditional cooperators is a substantial risk because they indiscriminately  
334 help defectors and allow their indirect invasion. We summarize mechanisms for  
335 preventing invasion by unconditional cooperators as follows:

336 **Rationality:** Not discriminating between cooperation and defection towards bad players  
337 when one-sided defection is individually rational, *i.e.*,  $T > 1$  (Fig. 2(a); feasible in SG  
338 and PD)

339 **Enforcement:** Unjustifying mutual cooperation with bad players (Fig. 2(b1, b2, d1, d2);  
340 feasible in SG, PD, and SH).

341 In previous works, the variants of the norms called standing and shunning employed the  
342 rationality mechanism, and those of the norm called Kandori employed the enforcement  
343 mechanism (see, *e.g.*, Refs. [10, 19]).

344 The social norms presented in Fig. 2(e) are successful in SH. They assign good reputa-  
345 tions to players that select cooperation (defection) towards good (bad) co-players. In other  
346 words, these norms possess an unfair bias for favouring reciprocators whereby they always  
347 regard reciprocation as a good deed. This property maintains mutual cooperation among  
348 reciprocators even when there is a non-negligible fraction of other strategists, and thus, it  
349 succeeds in enlarging their basin of attraction. The other features of these norms are that  
350 their punishments of defection against good players (*i.e.*, outcomes T and P) can be milder  
351 than those in case of SG and PD, and that they have the enforcement mechanism introduced  
352 above.

## 353 B. Information use and emerging uncontrollability of reputation

354 For determining a focal player's reputation, social norms in our model use three sources  
355 of information, *i.e.*, the focal player's action, his/her co-player's action, and the co-player's  
356 reputation (see Tab. III). The stability and efficiency of indirect reciprocity is generally sen-  
357 sitive to which kinds of information are available. Early studies of indirect reciprocity in  
358 evolutionary games focused on the so-called first-order assessment, which only takes into

359 account a focal player’s past action ( $a_{\text{focal}}$  in Tab. III) for determining the player’s reputa-  
 360 tion [5, 6, 40]. However, the first-order assessment is not sufficient to stabilize reciprocation  
 361 except adopting special assumptions [7–9, 38, 41]. Reciprocation can be stable when the  
 362 assessment uses at least two sources of information: a focal player’s action and his/her  
 363 co-player’s reputations ( $a_{\text{focal}}$  and  $r_{\text{co}}$  in Tab. III). This is because they enable one to dis-  
 364 tinguish naïve defection (*i.e.*, defection against a good player) and defection to be justified  
 365 (*i.e.*, defection against a bad player). There are a couple of reviews that explain the issue  
 366 of justified defection [33, 36, 42]. It should be noted that the justified defection is not an  
 367 only way to stabilize indirect reciprocity; *e.g.*, the ‘shunning’ social norm [12, 19, 43] and  
 368 the ‘tolerant scoring’ [22, 44].

369 The availability of information introduces not only the justified defection, but also a  
 370 ‘gamble’. In our model, players face with a gamble in which the player’s new reputation  
 371 may depend on the co-player’s action ( $a_{\text{co}}$  in Tab. III). This means that an actor in a game  
 372 cannot fully control his/her new reputation by taking an appropriate action. Such a sort of  
 373 uncontrollability has tacitly appeared in previous studies. For example, under the shunning  
 374 social norm, when an actor meets a bad recipient, the actor always receives a bad reputation  
 375 regardless of his/her behavior [36, 43]. In the shunning norm, an actor faces with a gamble  
 376 on what kind of recipient he/she encounters. This is also true in the simple-standing norm,  
 377 in which an actor always receives a good reputation when he/she by chance encounters a  
 378 bad recipient [12]. In contrast to such uncontrollability in encounters, our model contains  
 379 another uncontrollability in the co-player’s actions. On the uncontrollability in the co-  
 380 player’s actions, our results have shown that successful social norms have the following  
 381 characteristics:

- 382 1. In PD and SG (see Fig. 2(c)), the uncontrollability disappears when a player encounters  
 383 a good co-players; the player’s new reputation when the game outcome is R and S (T  
 384 and P) is consistently good (bad) regardless of the co-player’s action.
- 385 2. In SH (see Fig. 2(e)), the uncontrollability disappears when a player adopts recipro-  
 386 cation; selecting cooperation with a good co-player (outcomes R and S) or defection  
 387 against a bad co-player (outcomes T and P) is consistently assessed as good.
- 388 3. Otherwise (*i.e.*, in cases when encountering a bad player in PD and SG, or when  
 389 selecting cooperation (defection) toward a bad (good) co-player in SH), the gamble



390 can emerge.

391 Under a population of reciprocators using one of the successful social norms, players typically  
392 encounter good co-players and select cooperation. The uncontrollability in the co-player's  
393 actions disappears in such typical scenarios, while it remains in rare scenarios (*i.e.*, encoun-  
394 tering bad players, selecting defection against a good player, *etc.*). This is also true for the  
395 uncontrollability in encounters in the previous studies.

396 In sum, our study revealed that indirect reciprocity is sometimes feasible even under social  
397 norms that include the apparently unreasonable uncontrollability in which a player can not  
398 necessarily anticipate his/her reputation by taking an appropriate action; the reputation  
399 may depend on the co-player's choice. However, in most cases under successful social norms  
400 that enable stable reciprocation, such uncertain situations are rare.

### 401 C. Comparison with the leading eight

402 Indirect reciprocity is also stabilized when using information about the focal player's  
403 action, the focal player's reputation, and the co-player's reputation (see the '3rd-order'  
404 column in Tab. III), under the so-called leading eight social norms in the donation game  
405 regime [10, 11]. Because the information used in the classical model in Refs. [10, 11] and  
406 that in the present model are different (see Tab. III), we cannot directly compare these two  
407 models. However, if we regard the co-player's actions C and D ( $a_{co}$  in Tab. III) as the focal  
408 player's reputations G and B ( $r_{focal}$  in Tab. III), respectively, the information use in our  
409 model corresponds with that in the classical model, and the social norms in Fig. 2(c) just  
410 agree with the leading eight. Therefore, we want to compare these two models using the  
411 above correspondence. The classical model, in contrast to ours, assumed more cognitively  
412 powerful players that use their own reputation information for their action rule. To clarify  
413 the difference between the two models, we extended our basic model to allow such intelligent  
414 players, and we found that only six of the leading eight norms survived the equilibrium  
415 selection (see Appendix D).

416 In Tab. IV, we show the two norms among the leading eight that failed to stabilize  
417 reciprocation in our extended model. In the classical model, the two norms succeed in  
418 stabilizing reciprocation when paired with the so-called OR strategy, with which a player  
419 defects against a bad co-player only when the player has a good reputation. Consider a

420 game involving two bad players, both adopting OR strategy. In the game, a focal player  
421 and his/her co-player both select cooperation, because each of them has a bad reputation.  
422 However, since the co-player selects cooperation, the outcome of the game may be either  
423 R or T, and the outcome T when playing with a bad co-player results in the focal player's  
424 good reputation. Therefore, in the two norms, the focal player can enjoy a better payoff  
425 if he/she switches the action to defection when  $T > 1$ , which is satisfied in the donation  
426 game. On the other hand, under the corresponding two norms in the classical model, when  
427 both players have bad reputations, cooperation and defection are regarded as good and bad,  
428 respectively (see Tab. IV). Thus, the OR strategy players have no incentive to switch their  
429 actions in the same situation.

430 In sum, a slight difference in the manner to use information destabilizes OR strategy,  
431 which is stable in the classical indirect reciprocity model and appears only when assuming  
432 players to be more intelligent. Note that, we have only analyzed the stability of homogeneous  
433 populations of the variants of reciprocator including OR strategy; it could be possible that  
434 a mixture of OR strategists and some more defective strategists, *e.g.*, normal reciprocators,  
435 is evolutionarily stable.

#### 436 D. Difference from two previous works that studied other than the donation game

437 Kandori was the first to study simultaneous-move games in community enforcement using  
438 reputation information [4]. His model and ours are fundamentally different in the following  
439 ways: 1) He investigated random matching games between two populations of players (*e.g.*,  
440 games between lenders and borrowers), whereas we studied random matching games between  
441 players in one population. 2) In his model, any equilibrium can be stabilized by long-term  
442 punishment (called  $T$ -period punishment) or by damaging the group-level reputation of the  
443 violator's population (called contagious equilibrium; see also Ref. [21]). They are strong  
444 devices for punishing defectors. In contrast, our model was not restricted to such strong  
445 punishments. We found that milder devices for punishment, the two mechanisms introduced  
446 above, are sufficient.

447 Uchida studied a Snowdrift-type donation game model [18]. He conducted a complete  
448 search on the entire combinations of third-order social norms and action rules and found  
449 that only two social norms, one, a variant of the Kandori norm and the other, called 'L4',

450 develop cooperative and stable populations of reciprocators. The ‘L4’ social norm damages  
451 reputation of players that defect against good players or cooperate with bad players when  
452 their own reputations are bad. If we regard a co-player’s cooperation and defection as the  
453 proxy of a focal player’s good and bad reputations, respectively, then the ‘L4’ norm implies  
454 that the outcomes T and P when a focal player plays with a good co-player or the outcome  
455 S when a focal player plays with a bad co-player are regarded as bad. Thus, the ‘L4’ norm  
456 is included in the norms shown in Fig. 2(a), which indeed are successful in SG. However, in  
457 the extended model that introduces more intelligent players, they are not successful in SG  
458 (see Tab. VI(a)).

#### 459 E. Limitations in the present study

460 In this study, we ignored the possibility that reputation is updated based on complete  
461 information about a focal player’s and his/her co-player’s actions and reputations, *i.e.*,  
462 fourth-order social norms (see Tab. III) [4, 11]. We also assumed that reputation information  
463 about a player is publicly shared among players, and ignored the possibility of nonpublic  
464 sharing in which players do not necessarily share reputation information, as studied in several  
465 previous works [9, 17, 20, 23]. These open questions should be explored in the future.

#### 466 ACKNOWLEDGMENTS

467 This work was supported by JSPS KAKENHI Grant No. 13J05595 to MN and JSPS  
468 KAKENHI Grant No. 25118006 to HO.

#### 469 Appendix A: Social norms in Fig. 2(e) satisfy the usefulness criterion in SH

470 We prove that CD players under the social norms shown in Fig. 2(e) have larger basins  
471 of attraction than CC players in competition with DD players in all the area of SH in the  
472 payoff space. Let  $\Delta f(x) \equiv f(b|x) - f(a|x)$  denote the payoff difference between the players  
473 adopting action rules  $b$  and  $a$  when the frequencies of  $a$ - and  $b$ -players are given by  $x$  and  
474  $1 - x$ , respectively. We assume that the action rules  $a$  and  $b$  are CD and DD, respectively.

475 By substituting  $a(G) = C$  and  $a(B) = b(G) = b(B) = D$  into Eq. (15), we obtain

$$\Delta f(x) = x [(q(G) - p(G))(S + T) + p(G)^2(S + T - 1)] - q(G)S. \quad (\text{A1})$$

476 Let  $x_{CC}^* = S/(S + T - 1)$  denote the critical fraction of CC players in competition with DD  
 477 players over which the CC players are advantageous than the DD players. If the basin of  
 478 attraction of CD players in competition with DD players is larger than that of CC players,  
 479 then  $\Delta f(x_{CC}^*) < 0$  holds true. The social norms in Fig. 2(e) imply in common that  $\phi(R, G) =$   
 480  $\phi(S, G) = \phi(T, B) = \phi(P, B) = 1 - \mu$  and  $\phi(R, B) = \mu$ . Substituting these  $\phi$  values,  
 481  $a(G) = C$ , and  $a(B) = b(G) = b(B) = D$  into Eq. (14), we solve  $\dot{p}(G) = \dot{q}(G) = 0$ , and  
 482 obtain

$$\begin{cases} p(G) &= 1 - \mu, \\ q(G) &= \frac{(1 - \mu) [1 - x(1 - \mu - \phi(P, G))]}{2 - \mu - x(1 - \mu) [1 - \phi(P, G) + \phi(T, G)] - (1 - x)\phi(P, G)}. \end{cases} \quad (\text{A2})$$

483 Substituting Eq. (A2) into Eq. (A1) at  $x = x_{CC}^*$ , we see that in an error-free limit (*i.e.*,  
 484  $\mu \rightarrow 0$ ),

$$\Delta f(x_{CC}^*) = -x_{CC}^* \frac{(1 - T) [1 - \phi_0(P, G)] - S [1 - \phi_0(T, G)]}{(1 - T) [2 - \phi_0(P, G)] - S [1 - \phi_0(T, G) + \phi_0(P, G)]} \quad (\text{A3})$$

485 holds true, where  $\phi_0(\cdot, \cdot) \in \{0, 1\}$  represents  $\phi(\cdot, \cdot) \in \{\mu, 1 - \mu\}$  in the error-free limit. Since  
 486  $1 > T > 0 > S$  holds true in SH, the denominator of the first term in the right-hand side  
 487 of Eq. (A3) is clearly positive. In order to let  $\Delta f(x_{CC}^*)$  be negative, the numerator, *i.e.*,  
 488  $(1 - T) [1 - \phi_0(P, G)] - S [1 - \phi_0(T, G)]$ , needs to be positive. This is satisfied unless both  
 489  $\phi_0(T, G)$  and  $\phi_0(P, G)$  are equal to 1, *i.e.*, when at least one pivot is B in the assessments  
 490 of defection (outcomes T and P) against good players, which indeed is met in the six norms  
 491 in Fig. 2(e).

## 492 **Appendix B: Accurate conditions for the stability of reciprocators**

493 Table V lists the accurate conditions for the stability of CD players under the social  
 494 norms shown in Fig. 2, which are derived from Eq. (13).

495 In Tab. V and hereafter, we denote a social norm in line as  $r_{11}r_{21}r_{31}r_{41}r_{12}r_{22}r_{32}r_{42}$ , where  
 496  $r_{ij}$  is either G, B, or ‘\*’ in row  $i$  and column  $j$  of the  $4 \times 2$  table that represents a norm as  
 497 seen in Fig. 2.

499 To stabilize reciprocation in SH, in Sec. VA, we mentioned that there is no need to  
 500 damage the reputation of defectors, since  $T < 1$  holds true in SH. To satisfy the usefulness  
 501 criterion, however, we need to do so.

502 Here we consider that in a population under the social norms in Fig. 2(e), the fraction  $x$   
 503 of players are reciprocators (*i.e.*, CD players) and the rest  $1 - x$  are unconditional defectors  
 504 (*i.e.*, DD players). A DD player, which has a good (bad) reputation with probability  $q(G)$   
 505 ( $q(B)$ ), encounters a CD player with probability  $x$  and achieves either of the outcomes T and  
 506 P with probabilities  $q(G)$  and  $q(B)$ , respectively. On the other hand, he/she encounters a  
 507 DD player with probability  $1 - x$  and here achieves the outcome P only. Thus, the expected  
 508 payoff of the DD player is  $x[q(G) \cdot T + q(B) \cdot 0] + (1 - x) \cdot 0 = xq(G)T$ . In a similar manner, the  
 509 expected payoff of a CD player is given by  $x \cdot 1 + (1 - x)[q(G) \cdot S + q(B) \cdot 0] = x + (1 - x)q(G)S$ ,  
 510 where it should be noted that a pair of CD players always achieve mutual cooperation (*i.e.*,  
 511 outcome R) because their reputations are always good under those norms. Therefore, the  
 512 payoff difference between the DD and CD players is

$$\Delta f = xq(G)T - [x + (1 - x)q(G)S] = x[q(G)(S + T) - 1] - q(G)S \quad (\text{C1})$$

513 in an error-free limit (*i.e.*,  $\mu \rightarrow 0$ ). By solving  $\Delta f = 0$ , we obtain the critical fraction of CD  
 514 players over which they are advantageous than DD players, given by

$$x_{\text{CD}}^* = \frac{q(G)S}{q(G)(S + T) - 1}. \quad (\text{C2})$$

515 On the other hand, the corresponding critical fraction of CC players is given by

$$x_{\text{CC}}^* = \frac{S}{S + T - 1}. \quad (\text{C3})$$

516 If the basin of attraction of CD players is larger than that of CC players in competition with  
 517 DD players,  $x_{\text{CD}}^* < x_{\text{CC}}^*$  holds true. This yields the condition for satisfying the criterion of  
 518 usefulness,

$$q(B) > 0. \quad (\text{C4})$$

519 The condition (C4) implies that at least we need to slightly reduce the reputation of DD  
 520 players for securing better robustness of CD players than that of CC players. Indeed,  
 521 although defection against a good player (*i.e.*, outcomes T and P) can be allowed in SH under  
 522 the norms in Fig. 2(e) (see Sec. V), these norms do not completely allow such defections.

524 **1. The extended model**

525 Here we assume that players are more intelligent; a player performs an action based on  
 526 his/her own as well as his/her co-player's reputation. In this case, an action rule is extended  
 527 as  $a(r_{\text{focal}}, r_{\text{co}})$ , where  $r_{\text{focal}}$  is the focal player's and  $r_{\text{co}}$  is the co-player's reputations. The  
 528 number of possible action rules is  $2^{2 \times 2} = 16$ . We denote the extended action rule in line  
 529 by  $s_{GG} s_{GB} s_{BG} s_{BB}$ , where  $s_{uv} = a(u, v) \in \{C, D\}$ . For example, the action rule CDCD  
 530 represents a normal reciprocator that selects cooperation and defection when his/her co-  
 531 player's reputation is good and bad, respectively, irrespective of his/her own reputation.

532 We are interested in identifying successful pairs of reciprocating action rules and social  
 533 norms that satisfy the criteria introduced in Sec. III. Among the 16 possible action rules,  
 534 we consider that four action rules, CDCC, CDCD, CDDC, and CDDD, are the variants of  
 535 reciprocator, because they perform reciprocation when they are of good reputation. There-  
 536 fore, the number of pairs to be examined is  $4 \times 256 = 1024$ . We replace all the action-rule  
 537 terms in Sec. III by the above extended ones, *e.g.*,  $a(r_{\text{co}}) \rightarrow a(r_{\text{focal}}, r_{\text{co}})$ , and perform the  
 538 same procedure except for the following three points.

539 **Change in the goodness criterion:** If players adopt action rules other than CDCD, the  
 540 fraction of good players does not necessarily agree with the frequency of cooperation; it is  
 541 given by

$$p(C) = \sum \sum p(r_{\text{focal}}) p(r_{\text{co}}) \mathbb{1}_C(a(r_{\text{focal}}, r_{\text{co}})), \quad (\text{D1})$$

where  $\mathbb{1}_C(\cdot)$  is an indicator function by which  $\mathbb{1}_C(C) = 1$  and  $\mathbb{1}_C(D) = 0$ . We redefine that a pair of an action rule and a social norm satisfies the criterion of goodness if

$$p(C) = 1 - O(\mu) \quad (\text{D2a})$$

and

$$\lim_{\mu \rightarrow 0} p(G) > \frac{1}{2} \quad (\text{D2b})$$

542 holds true. Note that the condition (D2b) is necessary in order to rule out possible pairs  
 543 of the CDDC action rule and some social norms whereby a majority of players are of bad  
 544 reputation but cooperative. In such a population, the CDDC players achieve mutual coop-

545 eration because they have bad reputations and thereby help bad players; here the symbols  
 546 G and B actually stand for ‘bad’ and ‘good’, respectively [10].

547 **Change in the stability criterion:** In the extended model, if a pair of an action rule and  
 548 a social norm satisfies the goodness criterion, the payoff difference between the mutants and  
 549 residents, *i.e.*,  $\Delta f$  in Eq. (11), is indeed at most of  $O(\mu^2)$ . Thus, we expand  $\Delta f$  by  $\mu$  as

$$\Delta f = d_0 + \mu d_1 + \mu^2 d_2, \quad (\text{D3})$$

550 and if

$$\begin{cases} d_0 < 0 & (d_0 \neq 0) \\ d_1 < 0 & (d_0 = 0 \text{ and } d_1 \neq 0) \\ d_2 < 0 & (d_0 = d_1 = 0 \text{ and } d_2 \neq 0) \end{cases} \quad (\text{D4})$$

551 holds true, we regard that the pair satisfies the criterion of stability.

552 **Change in the usefulness criterion:** In the extended model, a reputation dynamic in  
 553 a polymorphic population (*cf.*, Eq. (14)) has possibly multiple stable equilibria, and which  
 554 equilibrium to be reached depends on the initial states. Therefore, we assume that all  
 555 the players have good reputations in the beginning, and numerically obtain an equilibrium  
 556 reached from the initial state.

## 557 2. Results

558 We examined the 1024 pairs of the variants of reciprocator (either CDCC, CDCD, CDDC,  
 559 or CDDD) and social norms. Unfortunately, no pair survives the equilibrium selection when  
 560 we consider the entire payoff space, *i.e.*,  $0 < T$  and  $S < 1$  (see Fig. 1). However, mutual  
 561 cooperation is Pareto efficient only when  $S+T < 2$  holds true (see, *e.g.*, Ref. [34]). Narrowing  
 562 the region of interest in the payoff space by adding the constraint  $S + T < 2$ , we identified  
 563 the successful 27 pairs shown in Tab. VI.

564 The pairs shown in Tab. VI(a, b1, b2, c) are included in Fig. 2(a, b1, b2, d1). Paired with  
 565 the CDCD action rule, *i.e.*, the normal reciprocator, the three social norms in Tab. VI(a)  
 566 satisfy the goodness and stability criteria in PD; the three social norms in Tab. VI(b1,2)  
 567 satisfy the goodness and stability criteria in PD and SH, whereas only those in Tab. VI(b1)  
 568 satisfy the usefulness criterion for SH; the two social norms in Tab. VI(c) satisfy the good-  
 569 ness, stability, and usefulness criteria in SH. Figure 2 shows successful twelve social norms

570 in the basic model, whereas Tab. VI(a, b1, b2, c) shows only eight. The lacking four  
571 pairs are CDCD-GGBBGGGB in Fig. 2(a), CDCD-GBBBBGGGB in Fig. 2(b2), and CDCD-  
572 GGBGB\*GG in Fig. 2(d2). In the extended model, the CDCD players are invaded by more  
573 intelligent mutants under these four norms. Moreover, the pairs in Tab. VI(a,b1,b2) are no  
574 longer stable in SG.

575 The 14 pairs shown in Tab. VI(d,e) satisfy the criteria of goodness, stability, and useful-  
576 ness in SH. The five pairs shown in Tab. VI(f,g) satisfy the criteria of goodness and stability  
577 in SG. In these 19 pairs, the dominating action rules are CDDC or CDDD whereby a player  
578 selects defection against a good co-player when the focal player has a bad reputation, and  
579 the social norms have an assessment in common such that the outcome P is always regarded  
580 as good, irrespective of the co-player's reputation. This assessment is plausible for the two  
581 action rules. Consider that in a population of CDDC or CDDD players, a bad player is  
582 playing a game with a good co-player. Because they adopt the CDDC or CDDD action rule,  
583 both of them select defection, *i.e.*, the outcome is P, and they receive good reputations under  
584 those norms. Intuitively, a player that adopts either of the two action rules infers about the  
585 co-player's next action from his/her own reputation, and the player selects defection when  
586 he/she is of a bad reputation. Such inference is effective in SH, which requires coordination  
587 (*i.e.*, mutual cooperation or defection) between two players. In SG, players have an incentive  
588 to select an action that is different to the co-player's, *i.e.*, C with D or D against C, and  
589 this characteristic of anti-coordination tends to break the mutual cooperation. However, the  
590 social norms shown in Tab. VI(f,g) assign bad reputations to such outcomes, *i.e.*, outcomes  
591 S and T, when a focal player plays with a good co-player. This assessments change SG into  
592 a coordination problem, and therefore, the two action rules perform well.

- 
- 593 [1] Trivers, R. L., 1971 The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–37. (doi:  
594 10.1086/406755).
- 595 [2] Sugden, R., 1986 *The economics of rights, co-operation and welfare*. Oxford: Blackwell.
- 596 [3] Alexander, R. D., 1987 *The biology of moral systems*. New York: Aldine de Gruyter.
- 597 [4] Kandori, M., 1992 Social norms and community enforcement. *Rev. Econ. Stud.* **59**, 63–80.  
598 (doi:10.2307/2297925).



- 599 [5] Nowak, M. A. & Sigmund, K., 1998 Evolution of indirect reciprocity by image scoring. *Nature*  
600 **393**, 573–577. (doi:10.1038/31225).
- 601 [6] Nowak, M. A. & Sigmund, K., 1998 The dynamics of indirect reciprocity. *J. Theor. Biol.* **194**,  
602 561–574. (doi:10.1006/jtbi.1998.0775).
- 603 [7] Leimar, O. & Hammerstein, P., 2001 Evolution of cooperation through indirect reciprocity.  
604 *Proc. R. Soc. Lond. B* **268**, 745–753. (doi:10.1098/rspb.2000.1573).
- 605 [8] Panchanathan, K. & Boyd, R., 2003 A tale of two defectors: the importance of standing  
606 for evolution of indirect reciprocity. *J. Theor. Biol.* **224**, 115–126. (doi:10.1016/S0022-  
607 5193(03)00154-1).
- 608 [9] Brandt, H. & Sigmund, K., 2004 The logic of reprobation: assessment and action rules for  
609 indirect reciprocation. *J. Theor. Biol.* **231**, 475–486. (doi:10.1016/j.jtbi.2004.06.032).
- 610 [10] Ohtsuki, H. & Iwasa, Y., 2004 How should we define goodness?—reputation dynamics in indi-  
611 rect reciprocity. *J. Theor. Biol.* **231**, 107–120. (doi:10.1016/j.jtbi.2004.06.005).
- 612 [11] Ohtsuki, H. & Iwasa, Y., 2006 The leading eight: social norms that can maintain cooperation  
613 by indirect reciprocity. *J. Theor. Biol.* **239**, 435–444. (doi:10.1016/j.jtbi.2005.08.008).
- 614 [12] Ohtsuki, H. & Iwasa, Y., 2007 Global analyses of evolutionary dynamics and exhaustive search  
615 for social norms that maintain cooperation by reputation. *J. Theor. Biol.* **244**, 518–531. (doi:  
616 10.1016/j.jtbi.2006.08.018).
- 617 [13] Ohtsuki, H., Iwasa, Y. & Nowak, M. A., 2009 Indirect reciprocity provides only a narrow  
618 margin of efficiency for costly punishment. *Nature* **457**, 79–82. (doi:10.1038/nature07601).
- 619 [14] Pacheco, J. M., Santos, F. C. & Chalub, F. A. C. C., 2006 Stern-judging: a simple, successful  
620 norm which promotes cooperation under indirect reciprocity. *PLoS Comput. Biol.* **2**, e178.  
621 (doi:10.1371/journal.pcbi.0020178).
- 622 [15] Takahashi, N. & Mashima, R., 2006 The importance of subjectivity in perceptual er-  
623 rors on the emergence of indirect reciprocity. *J. Theor. Biol.* **243**, 418–436. (doi:  
624 10.1016/j.jtbi.2006.05.014).
- 625 [16] Uchida, S. & Sigmund, K., 2010 The competition of assessment rules for indirect reciprocity.  
626 *J. Theor. Biol.* **263**, 13–19. (doi:10.1016/j.jtbi.2009.11.013).
- 627 [17] Uchida, S., 2010 Effect of private information on indirect reciprocity. *Phys. Rev. E* **82**, 036111.  
628 (doi:10.1103/PhysRevE.82.036111).
- 629 [18] Uchida, S., 2011 Indirect reciprocity in snowdrift games. *Pacific Sci. Rev.* **12**, 190–195.

- 630 [19] Nakamura, M. & Masuda, N., 2011 Indirect reciprocity under incomplete observation. *PLoS*  
631 *Comput. Biol.* **7**, e1002113. (doi:10.1371/journal.pcbi.1002113).
- 632 [20] Nakamura, M. & Masuda, N., 2012 Groupwise information sharing promotes ingroup fa-  
633 voritism in indirect reciprocity. *BMC Evol. Biol.* **12**, 213. (doi:10.1186/1471-2148-12-213).
- 634 [21] Masuda, N., 2012 Ingroup favoritism and intergroup cooperation under indirect reciprocity  
635 based on group reputation. *J. Theor. Biol.* **311**, 8–18. (doi:10.1016/j.jtbi.2012.07.002).
- 636 [22] Tanabe, S., Suzuki, H. & Masuda, N., 2013 Indirect reciprocity with trinary reputations. *J.*  
637 *Theor. Biol.* **317**, 338–347. (doi:10.1016/j.jtbi.2012.10.031).
- 638 [23] Oishi, K., Shimada, T. & Ito, N., 2013 Group formation through indirect reciprocity. *Phys.*  
639 *Rev. E* **87**, 030801. (doi:10.1103/PhysRevE.87.030801).
- 640 [24] Wedekind, C. & Milinski, M., 2000 Cooperation through image scoring in humans. *Science*  
641 **288**, 850–852. (doi:10.1126/science.288.5467.850).
- 642 [25] Milinski, M., Semmann, D., Bakker, T. C. & Krambeck, H. J., 2001 Cooperation through  
643 indirect reciprocity: image scoring or standing strategy? *Proc. R. Soc. Lond. B* **268**, 2495–  
644 2501. (doi:10.1098/rspb.2001.1809).
- 645 [26] Bolton, G. E., Katok, E. & Ockenfels, A., 2005 Cooperation among strangers  
646 with limited information about reputation. *J. Public. Econ.* **89**, 1457–1468. (doi:  
647 10.1016/j.jpubeco.2004.03.008).
- 648 [27] Seinen, I. & Schram, A., 2006 Social status and group norms: indirect reciprocity in a repeated  
649 helping experiment. *Euro. Econ. Rev.* **50**, 581–602. (doi:10.1016/j.eurocorev.2004.10.005).
- 650 [28] Sommerfeld, R. D., Krambeck, H.-J., Semmann, D. & Milinski, M., 2007 Gossip as an alter-  
651 native for direct observation in games of indirect reciprocity. *Proc. Natl. Acad. Sci. U.S.A.*  
652 **104**, 17435–17440. (doi:10.1073/pnas.0704598104).
- 653 [29] Engelmann, D. & Fischbacher, U., 2009 Indirect reciprocity and strategic reputation  
654 building in an experimental helping game. *Games Econ. Behav.* **67**, 399–407. (doi:  
655 10.1016/j.geb.2008.12.006).
- 656 [30] Pfeiffer, T., Tran, L., Krumme, C. & Rand, D. G., 2012 The value of reputation. *J. R. Soc.*  
657 *Interface* **9**, 2791–2797. (doi:10.1098/rsif.2012.0332).
- 658 [31] Yoeli, E., Hoffman, M., Rand, D. G. & Nowak, M. A., 2013 Powering up with indirect reci-  
659 procity in a large-scale field experiment. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10424–10429.  
660 (doi:10.1073/pnas.1301210110).

- 661 [32] Greif, A., 1989 Reputation and coalitions in medieval trade: evidence on the Maghribi traders.  
662 *J. Econ. Hist.* **49**, 857–882. (doi:10.1017/S0022050700009475).
- 663 [33] Sigmund, K., 2010 *The calculus of selfishness*. Princeton: Princeton University Press.
- 664 [34] Macy, M. W. & Flache, A., 2002 Learning dynamics in social dilemmas. *Proc. Natl. Acad.*  
665 *Sci. U.S.A.* **99**, 7229–7236. (doi:10.1073/pnas.092080099).
- 666 [35] Santos, F. C., Pacheco, J. M. & Lenaerts, T., 2006 Evolutionary dynamics of social dilemmas  
667 in structured heterogeneous populations. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3490–3494. (doi:  
668 10.1073/pnas.0508201103).
- 669 [36] Nowak, M. A. & Sigmund, K., 2005 Evolution of indirect reciprocity. *Nature* **437**, 1291–1298.  
670 (doi:10.1038/nature04131).
- 671 [37] Sigmund, K., 2012 Moral assessment in indirect reciprocity. *J. Theor. Biol.* **299**, 25–30.  
672 (doi:10.1016/j.jtbi.2011.03.024).
- 673 [38] Fishman, M. A., 2003 Indirect reciprocity among imperfect individuals. *J. Theor. Biol.* **225**,  
674 285–292. (doi:10.1016/S0022-5193(03)00246-7).
- 675 [39] Weber, H.-J. & Arfken, G. B., 2003 *Essential Mathematical Methods for Physicists*. San Diego:  
676 Academic Press.
- 677 [40] Takahashi, N., 2000 The Emergence of Generalized Exchange. *Am. J. Sociol.* **105**, 1105–1134.
- 678 [41] Brandt, H. & Sigmund, K., 2005 Indirect reciprocity, image scoring, and moral hazard. *Proc.*  
679 *Natl. Acad. Sci. U.S.A.* **102**, 2666–2670. (doi:10.1073/pnas.0407370102).
- 680 [42] Brandt, H., Ohtsuki, H., Iwasa, Y. & Sigmund, K., 2007 A survey of indirect reciprocity. In  
681 *Math. Ecol. Environ. Sci.* (eds. Y. Takeuchi, Y. Iwasa & K. Sato), number 1971 in Biological  
682 and Medical Physics, Biomedical Engineering, pp. 21–49. Berlin, Heidelberg: Springer Berlin  
683 Heidelberg. (doi:10.1007/978-3-540-34428-5\\_3).
- 684 [43] Takahashi, N. & Mashima, R., 2003 The emergence of indirect reciprocity: Is the standing  
685 strategy the answer?
- 686 [44] Berger, U., 2011 Learning to cooperate via indirect reciprocity. *Games Econ. Behav.* **72**,  
687 30–37. (doi:10.1016/j.geb.2010.08.009).
- 688 [45] Suzuki, S. & Akiyama, E., 2005 Reputation and the evolution of cooperation in sizable groups.  
689 *Proc. R. Soc. Lond. B* **272**, 1373–1377. (doi:10.1098/rspb.2005.3072).
- 690 [46] Brandt, H. & Sigmund, K., 2006 The good, the bad and the discriminator—errors in direct  
691 and indirect reciprocity. *J. Theor. Biol.* **239**, 183–194. (doi:10.1016/j.jtbi.2005.08.045).

- 692 [47] Suzuki, S. & Akiyama, E., 2008 Evolutionary stability of first-order-information indirect reci-  
693 procity in sizable groups. *Theor. Popul. Biol.* **73**, 426–436. (doi:10.1016/j.tpb.2007.12.005).
- 694 [48] Masuda, N. & Nakamura, M., 2012 Coevolution of trustful buyers and cooperative sellers in  
695 the trust game. *PLoS One* **7**, e44169. (doi:10.1371/journal.pone.0044169).
- 696 [49] Panchanathan, K. & Boyd, R., 2004 Indirect reciprocity can stabilize cooperation without the  
697 second-order free rider problem. *Nature* **432**, 499–502. (doi:10.1038/nature02978).
- 698 [50] Suzuki, S. & Akiyama, E., 2007 Evolution of indirect reciprocity in groups of vari-  
699 ous sizes and comparison with direct reciprocity. *J. Theor. Biol.* **245**, 539–552. (doi:  
700 10.1016/j.jtbi.2006.11.002).
- 701 [51] Uchida, S. & Sasaki, T., 2013 Effect of assessment error and private information  
702 on stern-judging in indirect reciprocity. *Chaos, Sol. Frac.* **56**, 175–180. (doi:  
703 10.1016/j.chaos.2013.08.006).
- 704 [52] Chalub, F. A. C. C., Santos, F. C. & Pacheco, J. M., 2006 The evolution of norms. *J. Theor.*  
705 *Biol.* **241**, 233–240. (doi:10.1016/j.jtbi.2005.11.028).
- 706 [53] Panchanathan, K., 2011 Two wrongs don’t make a right: the initial viability of different  
707 assessment rules in the evolution of indirect reciprocity. *J. Theor. Biol.* **277**, 48–54. (doi:  
708 10.1016/j.jtbi.2011.02.009).
- 709 [54] Martinez-Vaquero, L. A. & Cuesta, J. A., 2013 Evolutionary stability and resistance to cheat-  
710 ing in an indirect reciprocity model based on reputation. *Phys. Rev. E* **87**, 052810. (doi:  
711 10.1103/PhysRevE.87.052810).

FIG. 1. Three types of social dilemmas. In the payoff space spanned by  $T$  and  $S$ , the game defined by the payoff matrix (1) is the Snowdrift game (SG) when  $T > 1 > S > 0$  (green region), the Prisoner's Dilemma game (PD) when  $T > 1 > 0 > S$  (red region), and the Stag Hunt game (SH) when  $1 > T > 0 > S$  (yellow region). The standard donation game is on the solid red line ( $S + T = 1$  ( $T > 1$ )). Schematic diagrams inside these regions represent dynamics in competitions between cooperators (C) and defectors (D). Arrows represent the direction of evolution. Solid and hollow circles represent stable and unstable rest points, respectively.

FIG. 2. Surviving social norms. The symbol '\*' indicates a placeholder to be replaced by two patterns: G or B. The symbol '†' indicates another placeholder to be replaced by three vertical patterns: G and B, B and G, or B and B. Thus, each table represents  $2^n \times 3^m$  norms where  $n$  and  $m$  are the numbers of '\*' and '†' in the table, respectively. The Venn diagram indicates stability in different social dilemmas: SG (green), PD (red), and SH (yellow). (a) Social norms that are stable in SG and PD. (b1,2) Norms that are stable in SG, PD, and SH. Only those in b1 satisfy the usefulness criterion for SH. (c) Common characteristics of a, b1, and b2 that are successful in SG and PD. (d1,2) Norms that are stable and meet the usefulness criterion for SH. (e) Common characteristics of b1, d1, and d2 that are successful in SH. Note that all the norms here satisfy the goodness criterion.

TABLE I. An example of a social norm. The rows represent outcomes of a game (R, S, T or P), the columns represent a co-player's reputations (G or B), and G and B in each pivot represent the reputations that a focal player receives.

	G	B
R	G	G
S	G	G
T	B	G
P	B	G

TABLE II. Meaning of symbols.

symbol	meaning
$a(r) \in \{C, D\}$	Resident player's action in response to his/her co-player that have reputation $r$ .
$b(r) \in \{C, D\}$	Mutant player's action in response to his/her co-player that have reputation $r$ .
$p(r)$	Fraction of resident players that have reputation $r$ .
$q(r)$	Fraction of mutant players that have reputation $r$ .
$g(u, v) \in \{R, S, T, P\}$	Outcome of a game when a focal player and his/her co-player select actions $u$ and $v$ , respectively.
$\phi(g, r) \in \{1 - \mu, \mu\}$	Probability that a focal player receives a good reputation when the outcome of the game is $g$ and his/her co-player has reputation $r$ .
$\psi(g) \in \{1, S, T, 0\}$	Payoff to a focal player when the outcome of the game is $g$ .

TABLE III. Information use in social norms. First-, second-, and third-order social norms have been studied previously. The columns of ‘information use’ indicate whether to use the information of a focal player’s action ( $a_{\text{focal}}$ ), a focal player’s reputation ( $r_{\text{focal}}$ ), a co-player’s action ( $a_{\text{co}}$ ), and/or a co-player’s reputation ( $r_{\text{co}}$ ). The column of ‘justified defection’ indicates the availability of justified defection. The columns of ‘uncontrollability’ indicate the possibility of the uncontrollability of reputation in encounters ( $r_{\text{co}}$ ) and in the co-player’s actions ( $a_{\text{co}}$ ).

norm class	information use				justified defection	uncontrollability		previous studies
	$a_{\text{focal}}$	$r_{\text{focal}}$	$a_{\text{co}}$	$r_{\text{co}}$		$r_{\text{co}}$	$a_{\text{co}}$	
1st-order	✓	-	-	-	-	-	-	Refs. [5, 6, 38, 41, 44–48]
2nd-order	✓	-	-	✓	✓	✓	-	Refs. [12, 13, 15–23, 49–51]
3rd-order	✓	✓	-	✓	✓	✓	-	Refs. [2, 4, 7–11, 14, 52–54]
our model	✓	-	✓	✓	✓	✓	✓	-
4th-order	✓	✓	✓	✓	?	?	?	-

TABLE IV. The two social norms among the leading eight, which failed to stabilize reciprocation in our extended model. The left and right tables show these corresponding norms in our model and in the classical model studied in Refs. [10, 11], respectively. The symbol ‘\*’ indicates a placeholder to be replaced by two patterns: G or B. In the right table, the columns (GG, GB, BG, or BB) indicate that a focal player and his/her co-player in a game have both good, good and bad, bad and good, or both bad reputations, respectively.

	G	B
R	G	*
S	G	G
T	B	G
P	B	B

 $\iff$ 

	GG	GB	BG	BB
C	G	*	G	G
D	B	G	B	B

TABLE V. Conditions for the stability of reciprocators under the social norms in Fig. 2.

panel	social norm	stability condition
(a)	GGBBGGGG	$(T = 1 \wedge S < 0) \vee T > 1$
	GGBBGGGB	$T > 1$
	GGBBGBGG	$(T = 1 \wedge S < 1/2) \vee T > 1$
	GGBBGBGB	$(T = 1 \wedge S < 0) \vee T > 1$
(b1,2)	GGBBB*G*	$S < T$
(d1,2)	GGGBB*GG	$(T < 3/2 \wedge S < T) \vee (T = 3/2 \wedge 1/6 < S < 3/2)$
	GGBGB*GG	$(T < 2 \wedge S < T) \vee (T = 2 \wedge 1/2 < S < 2)$



TABLE VI. Surviving social norms when assuming more intelligent players.

group	action rule - social norm	stability			usefulness
		SG	PD	SH	SH
(a)	CDCD - GGBBG*GG	-	✓	-	-
	CDCD - GGBBGBGB	-	✓	-	-
(b1)	CDCD - GG BBB*GG	-	✓	✓	✓
(b2)	CDCD - GG BBBBBGB	-	✓	✓	-
(c)	CDCD - GG GBB*GG	-	-	✓	✓
(d)	CDDC - GG BGG**G	-	-	✓	✓
	CDDC - GG BGB*BG	-	-	✓	✓
(e)	CDDD - GG BG***G	-	-	✓	✓
(f)	CDDC - GG BGGBBG	✓	-	✓	-
(g)	CDDD - GG BG*B*G	✓	-	✓	-