Voluntary Redistribution Mechanism in Asymmetric Coordination Games*

Masaki Aoyagi[†] Naoko Nishimura[‡] Yoshitaka Okano[§]

September 2, 2020

Abstract

An inequality game is an asymmetric 2×2 coordination game in which player 1 earns a substantially higher payoff than player 2 except in the inefficient Nash equilibrium (NE). The two players may have either common or conflicting interests over the two NE. This paper studies a redistribution scheme which allows the players to voluntarily transfer their payoffs after the play of an inequality game. We find that the redistribution scheme induces positive transfer from player 1 to player 2 in both common- and conflicting- interest games, and is particularly effective in increasing efficient coordination and reducing coordination failures in conflicting-interest games. We explain these findings by considering reciprocity by player 1 in response to the sacrifice made by player 2 in achieving efficient coordination in conflicting-interest games.

Key words: equity, efficiency, transfer, reciprocity, sacrifice.
Journal of Economic Literature Classification Numbers: C72, D31, D63

1 Introduction

Coordination failures are some of the most important sources of economic inefficiencies. Coordination games have been used extensively to study both theoretically and experimentally the sources and remedies of coordination failures. As observed by Crawford et al. (2008), one instance where severe coordination failures take place is when coordination entails asymmetric payoffs for the players involved. In this paper, we consider a class of 2×2 coordination games with highly asymmetric payoffs

*We are very grateful to the Editor and a referee for the comments that led to a significant improvement of the paper. Financial support from the JSPS (grant numbers: 22330061, 23530216, 24330064, 24653048, 15K13006, 15H03328, 16H03597, 16K17088) and the Joint Usage/Research Center at ISER, Osaka University, and the International Joint Research Promotion Program of Osaka University is gratefully acknowledged.

[†]Osaka University, email:aoyagi@iser.osaka-u.ac.jp.
[‡]Ritsumeikan University, email:nnaoko@fc.ritsumei.ac.jp.
[§]Kochi University of Technology, email:okano.yoshitaka@kochi-tech.ac.jp.

between the players, and examine through laboratory experiments if ex post voluntary transfer of payoffs helps eliminate coordination failures and increase efficient coordination.

Formally, an *inequality game* is a 2×2 coordination game with two Nash equilibria (NE) (X, X) and (Y, Y). Player 1 earns a strictly higher payoff than player 2 at every action profile except at (Y, Y), where they earn the same payoff. However, the sum of payoffs at (Y, Y) is substantially lower than that at (X, X), implying a tension between efficiency and equity. The inequality games are further classified into **C**om**M**on-interest (CM) inequality games in which both players' payoffs are higher at (X, X) than at (Y, Y), and **C**on**F**licting-interest (CF) inequality games in which player 1's payoff is higher but player 2's payoff is lower, at (X, X) than at (Y, Y). Examples of CM and CF inequality games are presented in Table 1. The inequality games are also parametrized by the degree of inequality between the two players' payoffs.

Table 1: Inequality games

	CM			CF	
	X	Y		X	Y
X	440, 110	60, 50	X	320, 80	60, 20
Y	380, 60	100, 100	Y	260, 60	100, 100

(X, X) efficient coordination; (Y, Y) equitable coordination.

The redistribution scheme we propose allows both players to voluntarily transfer their payoffs to the other player after the play of the inequality game. Although such a scheme will have no impact on the outcome of the game under self-interested preferences, our main objective is to analyze its functioning in a laboratory where subjects' motivation may come from other sources than self-interest.

The efficiency-equity trade-off between the two NE in the inequality games can be a fundamental source of coordination failures. The players playing CM and CF games also can have sufficiently different motivations when choosing their actions. In CM games, coordination on the Pareto efficient profile (X, X) will result unless the players are, or expect the other player to be, sufficiently inequality averse. In CF games, on the other hand, coordination on (X, X) would be more difficult since it entails a material sacrifice by player 2 compared with coordination on (Y, Y). With ex post redistribution, this difference in the motivations between the CM and CF games can have a significant impact on the final outcome. In other words, when (X, X) is realized in the CM games, player 1 may interpret it as a result of player 2's self-interested behavior, and may find little reason to reciprocate 2's choice of X with payoff transfer to him. On the other hand, if (X, X) is realized in CF games, player 1 may interpret it as resulting from 2's *self-sacrifice* to achieve an outcome which benefits 1. Player 1 may hence have incentive to reciprocate this with payoff transfer. Expecting this, however, player 2 may strategically choose X in CF to his own benefit.

In our experiments, each subject is randomly assigned the role of either player 1 or player 2, and is randomly and anonymously matched with a subject who is assigned the other role. The experiment consists of three parts with the subject role fixed throughout. In the first part, we have a half of the subjects in each role make a dictator decision over the action profiles of each inequality game. In the second part, the subjects play a series of inequality games in a standard way. In the third part, they play the inequality games under the redistribution scheme. Our design choice to have the same set of subjects play the games with no redistribution first and then with redistribution next, and provide in the instructions detailed information on how the payoffs are determined by the actions, is motivated by the importance of having the subjects understand the externalities involved in their decision making in the inequality games and the consequences of possible coordination failures. The within-subject design also allows us to associate the heterogeneity in the subjects' behavior with the difference in their types which are likely linked to their preferences and beliefs about the behavior of the other player.

Our results show that the redistribution scheme induces significantly positive transfer by player 1 in both CM and CF games. Positive transfer takes place almost exclusively when player 2 chooses action X which corresponds to the efficient NE preferred by player 1. The size and frequency of transfer is higher in CF games than in CM games, and increasing inequality increases the size of transfer but not the frequency of positive transfer. Comparison of the results with and without the redistribution scheme shows that the scheme induces the efficient NE (X, X) strongly significantly in CF games, but only weakly in CM games. We also find that the scheme increases the sum of the two players' payoffs significantly in CF games but only insignificantly in CM games. Interestingly, however, the scheme has on average little impact on player 1's payoff in both games, and the increase in efficiency comes almost entirely from the increase in player 2's payoff. On the other hand, equity as measured by the payoff ratio between the two players is significantly improved by the redistribution scheme in both CF and CM games.

Since the introduction of ex post payoff redistribution has no impact on the behavior of self-interested individuals, the observed increase in efficient coordination and positive transfer imply the presence of distributive social preferences and/or reciprocity. We attempt to identify the source of these effects based on some key observations. In particular, we remark that positive transfer by player 1 to player 2 takes place almost exclusively following 2's choice of X. This suggests that player 1 reciprocates player 2's action choice that benefits player 1. Furthermore, the observed difference between CM and CF games suggests that player 1 perceives the level of kindness entailed in 2's choice of X differently in the two games. Specifically, the choice of X by player 2 can result from self-interest in CM, but entails a sacrifice in CF. We suppose that player 1's reciprocity is strengthened by the presence of self-sacrifice by player 2, and postulate a psychological utility function that explicitly accounts for sacrifice. Taking advantage of the within-subject design, we also attempt to identify the subjects' motivations by examining their behavior in different tasks. In particular, we find that the increased choice of action X by the role 2 subjects in the redistribution scheme is likely motivated by self-interest: They choose X in anticipation of the choice of X and positive transfer by role 1.

The literature discusses social preferences almost exclusively in setups such as public good games, prisoners' dilemma games, ultimatum games and trust games. Behind the lack of analysis of social preferences in coordination games may be the intuitive perception that social preferences will only contribute to an increase in coordination. Our use of coordination games with payoff inequality and a tension between equality and efficiency presents a formal framework to test this intuition. Our findings suggest that whether social preferences contribute to an increase in coordination depends on the specification of payoffs as well as the existence of a reciprocation opportunity. As mentioned above, we find sacrifice made by one player to be a strong inducement to positive reciprocity from the other player.

The paper is organized as follows. The next section discusses the related literature. The inequality game and the redistribution scheme are described in Section 3. Section 4 describes the experimental design, and Section 5 presents the analysis. The motive behind the observed action choices and transfer decisions is discussed in Section 6. We conclude in Section 7.

2 Related Literature

Reciprocity-based mechanisms originate in the literature on public good games. Reciprocity in the form of a punishment or disapproval of other players is the focus of early study by Fehr and Gächter (2000) and Masclet et al. (2003).[1] While reciprocity is at the core of our analysis, asymmetry between the players in our model offers a significantly different environment from the symmetric environment in the early literature.

The public good literature also offers extensive research on the possible distortion of behavior associated with inequality among the players: Asymmetry is introduced either in the level of individual return (MPCR - marginal per capita return), or in the level of initial endowment (income) of each individual. The findings are largely inconclusive.[2] Combining asymmetry with redistribution in public good games, Dekel

[1]Andreoni et al. (2003) find that the punishment option has a much stronger impact on the proposer's behavior than the reward option in a dictator-like giving game. Fehr and Rockenbach (2003) however show that the intention of imposing a sanction can induce a non-cooperative behavior in the trust game. Houser et al. (2008) examines whether intention to sanction is more important than the mere presence of sanctions.

[2]Buckley and Croson (2006) find that the low-income subjects give a higher percentage of their income to the public good than the high-income subjects, whereas Hofmeyr et al. (2007) observe no

et al. (2017) and Gangadharan et al. (2017) present analysis most closely related to the present paper. When players with positive and negative MPCR's interact, and redistribution takes the form of either a punishment or reward, Dekel et al. (2017) observe that communication coupled with a reward increases contribution substantially. When players may ex post reward the others, Gangadharan et al. (2017) also find a positive impact of communication on both earnings and contribution, but show that its impact is significantly weakened in the presence of heterogeneity in MPCR.[3] While the present model shares many features with the papers on public good games with heterogeneity and ex post redistribution, its use of coordination games highlights the role of reciprocity more clearly. Specifically, it is intuitive that player 2's choice of X corresponding to the efficient coordination is a favor given to player 1, and that payoff transfer from 1 to 2 is a direct way of returning the favor.[4,5]

Turning to the extensive literature on coordination games experiments, the primary focus is on the comparison between payoff dominance and risk dominance as the effective predictor of the outcome of play.[6] The literature on coordination games also investigates ways to eliminate coordination failures, and finds mixed evidence on the effectiveness of forward induction and correlated equilibrium recommendations.[7] Our analysis avoids the comparison of risk- and payoff-dominance by using games that have a constant level of risk dominance for the efficient NE. More importantly, while forward induction or correlated equilibrium recommendations are based on self-regarding preferences, the working of the redistribution mechanism

impact of heterogeneity on the contribution. Oxoby and Spraggon (2013) find that heterogeneity significantly lowers contributions.

[3]Other public good experiments with redistribution include Uler (2011), who studies income redistribution under exogenous tax rates, and Belafoutas et al. (2013), who let the subjects choose the redistribution rate before the contribution decisions.

[4]Among other differences, our design does not involve explicit communication or repetition of the game between the same pair of subjects.

[5]Voluntary redistribution following a real-effort tournament is studied by Erkal et al. (2011), who study payoff transfer between the first-ranked and second-ranked subjects in the context of social preferences. See also Ohtake et al. (2013). Unlike in the present paper, however, this literature provides no analysis of the action choice in the first stage with or without the redistribution possibility.

[6]Cooper et al. (1990, 1992), Straub (1995), Van Huyck et al. (1990), Goeree and Holt (2005), among others, observe that risk dominance predicts subjects' play better than payoff dominance. Cachon and Camerer (1996) propose loss-avoidance as a selection principle.

[7]Cooper et al. (1993) and Evdokimov and Rustichini (2016) find support for forward induction in the battle of the sexes (BOS) game, whereas Huck and Müller (2005) suggest that the first-mover principle is important rather than forward induction. Among those who take the correlated equilibrium approach, Cason and Sharma (2007) observe a difference in subjects' behavior when they are matched against each other, and against a computer which always follows recommendations. Duffy and Feltovich (2010) find that the recommendations are followed more often when they are payoff-enhancing compared with the NE of the game. Bone et al. (2013) also find that the payoff specification affects subjects' obedience to recommendations. Anbarcı et al. (2018) find a negative impact of payoff asymmetry on obedience.

hinges on social preferences.

3 Models of Inequality and Redistribution

3.1 Inequality Games

Formally, an inequality game G is a 2×2 coordination game: Each player i chooses his action x_i from the set $\{0, 1\}$, and their payoff functions are given by

$$\begin{aligned} g_1(x) &= a(1-x_1)(1-x_2) + bx_1 + c_1x_2, \\ g_2(x) &= a(1-x_1)(1-x_2) + bx_2 + c_2x_1. \end{aligned} \tag{1}$$

For the interpretation of these payoff functions, suppose that each player i chooses whether to allocate his resource to either their private activity ($x_i = 1$) or an activity toward a public project ($x_i = 0$). The private activity generates positive externalities to the other player, whereas the public project results in a success if and only if both players allocate their resources to it. The successful public project is worth a to each player, whereas player i's private activity is worth b to himself and worth c_j to the other player j. When both players engage in private activities, the utility of each player is simply the sum of the benefits from his and the other player's activities.[8]

We suppose that the externality benefit that 2's private activity creates for 1 is larger than the externality benefit that 1's private activity creates for 2:

$$a > b > 0 \quad \text{and} \quad c_1 > c_2 > 0.$$

Note that the second condition is the only source of inequality between the two players. Writing X for $x_i = 1$, and Y for $x_i = 0$, we can depict the payoff table as in Table 2.[9]

Since $a > b > 0$, both (X, X) and (Y, Y) are pure NE. We also assume:

- $c_1 + c_2 > 2(a - b) \;\Leftrightarrow\; (X, X)$ uniquely maximizes the sum of payoffs.
- $c_1 > b > c_2 \;\Leftrightarrow\; g_1(x) > g_2(x)$ for $x \neq (Y, Y)$.
- $2b > a \;\Leftrightarrow\; (X, X)$ is *risk dominant.*

[8]The experimental instructions use neutral phrasing and express $1 - x_1 = M$ and $x_1 = N$.

[9]For the interpretation of the asymmetric payoffs, consider for example neighboring countries 1 and 2 that have environmental issues between them. $x_i = 1$ corresponds to the reduction of air pollution in country i, and $x_i = 0$ corresponds to the reduction of the pollution of public waters between them. Because of the dominant wind direction, reduction in air pollution in country 1 yields a relatively small benefit to country 2, but reduction in air pollution in country 2 yields a larger benefit to country 1 than it does to country 2 itself. On the other hand, water pollution cannot be reduced without the joint effort from the two countries. As another example, consider two workers who must allocate their effort between a production line and product development. Worker 1 is inexperienced whereas worker 2 is experienced. Product development requires joint effort from both workers. On the other hand, effort in the production line by either worker yields benefits to both of them with the spillover from the experienced worker to the inexperienced worker large and the spillover in the other direction small.

Table 2: Inequality game: $c_1 > c_2$

P1 \ P2	X		Y	
X	$b+c_1$	$b+c_2$	b	c_2
Y	c_1	b	a	a

It follows from the second condition that (Y, Y) is the only profile in which the two players earn the same payoff. We further restrict attention to the following subclasses of inequality games: An inequality game has **C**om**M**on-interest (CM) if $b + c_1 > b + c_2 > a$, and has **C**on**F**licting-interest (CF) if $b + c_1 > a > b + c_2$. In other words, if an inequality game has CM, then both players 1 and 2 prefer the NE (X, X) to the NE (Y, Y) (in terms of material payoffs), whereas if it has CF, then player 1 prefers (X, X) to (Y, Y) and player 2 prefers (Y, Y) to (X, X).

Our experiments use three CM inequality games denoted CM2, CM4 and CM6, and three CF inequality games denoted CF2, CF4 and CF6. The suffix represents the degree of inequality between the players and is equal to the payoff ratio at (X, X):

$$\frac{g_1(X,X)}{g_2(X,X)} = \frac{b+c_1}{b+c_2} = k \text{ in CF}k \text{ and CM}k.$$

We set our parameters as in Table 3. Since $g_1(X, X) - g_2(X, X) = c_1 - c_2$, within each class of games, the larger is k, the larger is the payoff difference at (X, X).[10] The resulting payoff tables are depicted in Tables 4 and 5. Note that all CM games are the same in terms of player 2's payoffs, and so are all CF games. Furthermore, since a and b are held constant in all games, so is the risk dominance level of (X, X).

Table 3: Parameter specifications

	CF2	CF4	CF6	CM2	CM4	CM6
a	100	100	100	100	100	100
b	60	60	60	60	60	60
c_1	100	260	420	160	380	600
c_2	20	20	20	50	50	50

[10]Inequality may be perceived in terms of the payoff difference rather than the payoff ratio. Our econometric analysis treats k as a dummy variable, and analyzes CM and CF games separately when examining the effect of k. Since the payoff difference also increases with the payoff ratio in each class of games, interpretation of inequality in terms of the payoff difference or payoff ratio is immaterial.

Table 4: CM inequality games

(a) CM2

	X		Y	
X	220,	110	60,	50
Y	160,	60	100,	100

(b) CM4

	X		Y	
X	440,	110	60,	50
Y	380,	60	100,	100

(c) CM6

	X		Y	
X	660,	110	60,	50
Y	600,	60	100,	100

Table 5: CF inequality games

(a) CF2

	X		Y	
X	160,	80	60,	20
Y	100,	60	100,	100

(b) CF4

	X		Y	
X	320,	80	60,	20
Y	260,	60	100,	100

(c) CF6

	X		Y	
X	480,	80	60,	20
Y	420,	60	100,	100

3.2 Voluntary Redistribution

Let u_i denote player i's final material payoff after the possible redistribution of their payoffs. Task 1 (T1) is the *baseline scheme* in which no redistribution takes place after the play of the inequality game G.

T1: The players' final payoffs equal their payoffs from G: $u_i = g_i$.

Task 2 (T2), on the other hand, is the *redistribution scheme* in which the players may give part or all of their payoffs to the other player after the play of the inequality game.

T2: After the players play G, they publicly observe the actions and payoffs, and then determine the amount of transfer to the other player. If player i gives $t_i \in [0, g_i]$ payoff points to player j ($i \neq j$), then i's final (material) payoff is given by ($t = (t_1, t_2)$)

$$u_i(x, t) = g_i(x) - t_i + t_j \quad \text{for } i = 1,\, 2,\, j \neq i. \tag{2}$$

Task 0 (T0) is the *dictator scheme* in which the final allocation is determined by only one of the players.

T0: Dictator Decision: One player in each pair makes a choice among four payoff pairs that correspond to the four cells of the payoff table.

3.3 Equilibrium under Reciprocity

Player i's strategy $x = (x_1, x_2) \in \{X, Y\}^2$ in T0 is the choice of an action profile, whereas his strategy in T1 is $x_i \in \{X, Y\}$. Player i's strategy in T2 is a pair

(x_i, σ_i), where $x_i \in \{X, Y\}$ is the action choice and $\sigma_i : \{X, Y\}^2 \to \boldsymbol{R}_+$ is the transfer function that determines transfer to the other player j for each realization of the action profile. The subgame perfect equilibrium (SPE) $(x, \sigma) = (x_i, \sigma_i)_{i=1,2}$ is defined in the standard manner.

The players have *self-interest* preferences if their utilities equal their material payoffs (2): $U_i \equiv u_i$ for $i = 1, 2$. Under self-interest preferences, no redistribution takes place in any SPE of T2 (*i.e.*, $\sigma_i(\cdot) \equiv 0$ for $i = 1, 2$), and x is consistent with an SPE of T2 if and only if it is a NE of G.

We say that the players have *reciprocity preferences* if they reward the other player through positive transfer for the favor given to them in the play of the inequality game. Specifically, we suppose that the reciprocity preferences are given by

$$U_i(x, t) = u_i(x, t) + \gamma_i(x) \log u_j(x, t), \tag{3}$$

where for $0 < \mu_i \le \nu_i$ ($i = 1, 2$),

$$\gamma_i(x) = \begin{cases} 0 & \text{if } g_i(x) \le a, \\ \mu_i & \text{if } g_i(x) > a \text{ and } g_j(x) \ge a, \\ \nu_i & \text{if } g_i(x) > a \text{ and } g_j(x) < a. \end{cases}$$

The second term of U_i represents player i's reciprocity concerns, and $\gamma_i(x)$ is the reciprocity weight that measures how kind j is toward i through his action choice in G. Specifically, player i takes $g_i(Y, Y) = a$ as the reference point, and considers j to be kind when j's alternative action choice $x_j = X$ raises i's payoff above a. That is, player i places a strictly positive weight $\gamma_i(x)$ on j's material payoff if and only if $g_i(x) > a$.[11] If j's choice of X not only raises i's payoff above a but also lowers j's own payoff from a, then i regards it as the sacrifice made by j in raising i's payoff, and rewards j even more strongly by placing a higher weight on j's material payoff. In the CM inequality games, for example, $\gamma_1(X, X) = \mu_1$ and $\gamma_2(X, X) = \mu_2$ since both players are better off at (X, X) than at (Y, Y). On the other hand, in the CF inequality games, $\gamma_1(X, X) = \nu_1$ and $\gamma_2(X, X) = 0$ since player 1 is better off and player 2 is worse off at (X, X) than at (Y, Y).[12] The following proposition holds for the SPE of T2 under reciprocal preferences.

[11] In order to obtain an interior solution in the optimal transfer choice, we use the log transformation of j's material payoff in the definition of U_i. Any concave transformation yields a qualitatively similar conclusion.

[12] The formulation of reciprocity in (3) closely corresponds to those in the literature based on the psychological game approach. In Rabin's formulation (Rabin, 1993) of reciprocity in simultaneous-move games, for example, player i's preferences are given by: $U_i(x_i) = u_i(x_i) + \tilde{f}_j \hat{f}_i(x_i)$, where u_i is i's material payoff, $\tilde{f}_j$ represents i's belief about how kind j is, and $\hat{f}_i(x_i)$ is j's payoff when i chooses action x_i given his belief about j's action. In T2, player j's kindness (in G) is revealed through his action choice in G, and hence i's beliefs need to play no role in the redistribution stage. See Dufwenberg and Kirchsteiger (2004). One important departure from the literature is our assumption that the reciprocity weight depends on whether or not there is a sacrifice by the

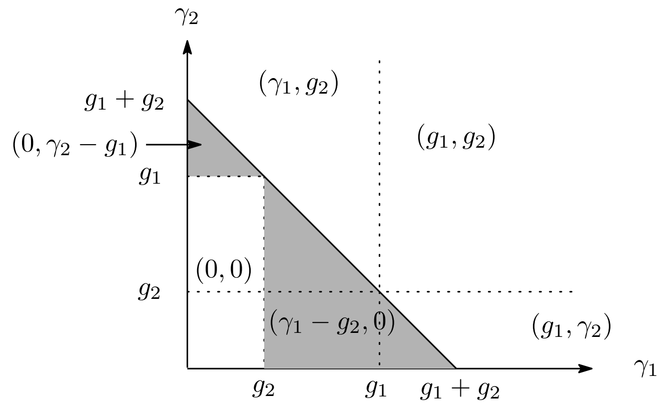


Figure 1: SPE transfer $t = (t_1, t_2)$ as a function of the reciprocity weights (γ_1, γ_2).

Proposition 1 *(SPE under reciprocity) Suppose that the players' preferences are given by (3).* (x, σ) *is an SPE of T2 if and only if* σ *is given by* ($i = 1$, *2,* $j \neq i$)

$$\sigma_i(x) = \begin{cases} \max\{\gamma_i(x) - g_j(x),\, 0\} & \text{if } \gamma_1(x) + \gamma_2(x) < g_1(x) + g_2(x), \\ \min\{\gamma_i(x),\, g_i(x)\} & \text{if } \gamma_1(x) + \gamma_2(x) > g_1(x) + g_2(x). \end{cases} \tag{4}$$

The SPE transfer $(\sigma_1(x), \sigma_2(x)) \equiv (t_1, t_2)$ in (4) is illustrated in Figure 1. In what follows, we restrict attention to the case where the reciprocity weights γ_i are not too large so that neither player transfers his entire payoff g_i.[13] When the parameters are in such a range, Figure 1 shows that at most one player makes a positive transfer, and which player does so depends on the relative magnitude of the reciprocity weights. Given that his payoff g_1 is much larger than player 2's payoff g_2, the figure shows that it is most likely player 1 who makes positive transfer. We can establish the following facts concerning an SPE (x, σ) of T2 under reciprocity preferences.

1. If a player makes positive transfer after (X, X) for some degree of inequality k, then he makes positive transfer for any larger inequality $k' > k$.

2. Player 1 makes positive transfer at (X, X) in CF but not in CM if $\mu_1 \leq b + c_2 < \nu_1$.

---

other player. Another preference specification that induces positive transfers is guilt-aversion as formulated by Charness and Dufwenberg (2006): A player feels guilty for not making a transfer when the other player expects it. In the present game, however, we find it difficult to specify plausible expectations.

[13]Specifically, we assume that $\gamma_1 + \gamma_2 < g_1 + g_2$ as in the first line of (4). At $x = (X, X)$, this is equivalent to $\mu_1 + \mu_2 \leq 2b + c_1 + c_2$ in CM and $\nu_1 \leq 2b + c_1 + c_2$ in CF. At $x = (Y, X)$, the equivalent condition is $\nu_1 \leq b + c_1$ in both CM and CF.

3. The set of SPE action profiles x is given by $\begin{cases} \{(X,X)\} & \text{if } \nu_1 > a, \\ \{(X,X),(Y,Y)\} & \text{if } \nu_1 < a. \end{cases}$

The reciprocity preferences as defined in (3) generate behavior different from self-interest only in T2, and no difference is expected either in T0 or T1. As alternative hypotheses, we consider the distributional social preferences as follows. The players have *inefficiency aversion* (IEA) preferences if they are concerned about the efficiency of an outcome as measured by the sum of their material payoffs. Specifically, we suppose that

$$U_i(x,t) = u_i(x,t) + \kappa_i\{u_1(x,t) + u_2(x,t)\} \quad \text{for } i = 1,\, 2, \tag{5}$$

where κ_i represents the degree of the inefficiency concerns relative to the own material payoff. The players have *inequality aversion* (IQA) preferences if they dislike inequality in their material payoffs. Specifically, we suppose that

$$U_i(x,t) = u_i(x,t) - \lambda_i\,|u_i(x,t) - u_j(x,t)| \quad \text{for } i = 1,\, 2,\, j \neq i, \tag{6}$$

where $\lambda_i > 0$ represents the degree of the inequality concerns relative to the own material payoff.[14] The implications of these preferences are as follows.[15]

4. In equilibrium under IEA, (X,X) is chosen more often as inequality increases, and the action profiles are the same under T1 and T2. No transfer takes place.

5. In equilibrium under IQA, the players choose (Y,Y) more often in T1 as inequality increases. If $\lambda_1 < \frac{1}{2}$, then no transfer takes place in T2 and (Y,Y) is chosen more often as inequality increases. If $\lambda_1 > \frac{1}{2}$, then (X,X) is chosen more often in T2 as inequality increases, and player 1 makes positive transfer except at (Y,Y).

4 Experimental Design

The experiments were conducted at the Experimental Economics Laboratory at the ISER, Osaka University, with the subjects recruited from undergraduate and graduate students of Osaka University of various majors. There were six sessions with a total of 124 subjects (four sessions of 20 subjects and two sessions of 22 subjects). No subject attended more than one session. The subjects in each session were divided randomly into two groups of the same size with the first group of subjects assigned the role of player 1, and the second group assigned the role of player 2. The player roles stay the same throughout the session. The role assignment

[14] For simplicity, this formulation defines inequality in terms of the payoff difference. A definition based on the payoff ratio is possible.

[15] See Appendix A.3 for the exact descriptions.

is done privately on the PC screen in front of each subject. The instruction presents the payoff formula (1), and provides its illustration by means of numerical examples and graphs.[16,17] The inclusion of the payoff formula is intended to help the subjects understand the source of inequality between the two roles, and also the externalities involved in their decision making. The payoff matrix is shown also on the PC screen in front of each subject. At the end of each session, the earning of a subject is computed from the sum of his/her payoff points during the session with the conversion rate of 1 payoff point to JPY1.3. The average earnings are JPY9946.1 for the role 1 subjects and JPY3122.8 for the role 2 subjects.[18] The subjects were also given a record sheet in which they describe their action and transfer choices as well as the reason behind those choices.

The experiments adopt the within-subject design and every session is divided into three *task blocks* that correspond to T0-T2 described in the previous section. Each task block in turn consists of six *rounds* that correspond to the six inequality games CF2-CF6 and CM2-CM6. In all four sessions, the ordering of the task blocks is fixed and given by

$$\text{T0} \to \text{T1} \to \text{T2}.$$

As mentioned in the Introduction, the fixed task order was adopted so that the subjects would become fully aware of the externalities involved in their decision making through the standard play of the inequality games in T1.[19] The six games appear in random order in the six rounds of each task block. In T0, a half of the role 1 subjects and a half of the role 2 subjects are randomly chosen to make a choice.[20] After every round, each subject observes his and the other player's action choices on their PC screen, and then is anonymously rematched to another subject of the

[16]The program was coded using z-tree (Fischbacher, 2007). The formula was included in the instructions for T0 and T1, and the graphs were included in the instructions for T1 which involved strategic interactions for the first time. The instructions for T2 didn't include them to avoid redundancy, but instead stated that the payoffs were determined by the same way as in T1. Apart from the six sessions with the payoff formula (1) in the instructions, we also had five sessions in which no payoff formula was presented in the instructions. See Section 6 for some discussion and Appendix A.2 for the analysis of these sessions. Every session had one additional task T3 which is not discussed in this paper: T3 involves the pre-play communication stage in which the two players simultaneously agree or disagree to have the voluntary post-play redistribution stage. The redistribution stage follows the game if and only if they both agree. (Communication here is hence unlike the free-form communication in Dekel et al. (2017) and Gangadharan et al. (2017).) T3 was given at the end of each session and didn't influence the subjects' behavior in the preceding tasks.

[17]See Appendix A.4 for an English translation of the instructions. The average time spent on tasks T0-T2 in six sessions is 89 minutes including time spent on instructions (20 minutes at the start, and 10 minutes between tasks). The subjects were given three minutes before each task to self-check their comprehension and ask questions.

[18]These translate to about US$90-127 for role 1 and US$28-40 for role 2 according to the exchange rates at the time of the experiments.

[19]This design choice is also based on the four pilot sessions which rotated task orders. See Footnote 39.

[20]Once chosen, they make choices in all six rounds of the T0 block (against different opponents).

opposite role in the stranger format.[21]

5 Analysis

5.1 Dictator decisions

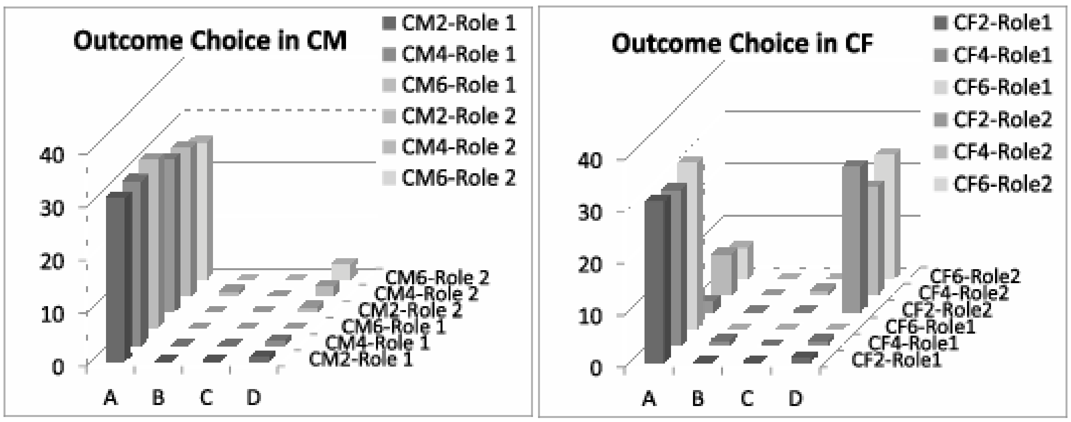


Figure 2: Outcome distribution in T0

We begin by examining the outcome of the dictator decision task (T0). Figure 2 shows the frequency of each of four choices for CM and CF games, where $A = (X, X)$, $B = (X, Y)$, $C = (Y, X)$, and $D = (Y, Y)$. As seen, the subjects' choices are almost entirely limited to (X, X) and (Y, Y), which are NE of the game.[22] Additionally, the role 1 subjects choose (X, X) more than 94% of the time, whereas the choice of the role 2 subjects is approximately reversed depending on whether the game is CM or CF. In CM, they choose (X, X) more than 90% of the time, whereas in most cases of CF, they choose (Y, Y) 80% of the time.[23] The inequality dummy k is mostly insignificant in both CM and CF.[24]

Role 1's choice of (X, X) is consistent with self-interest and IEA, whereas role 2's choice of (X, X) in CM and (Y, Y) in CF is consistent with self-interest and IQA.

[21] Since ten or eleven pairs are formed in each session which consists of three task blocks of six rounds each, the probability that a subject is matched with the same opponent is positive.

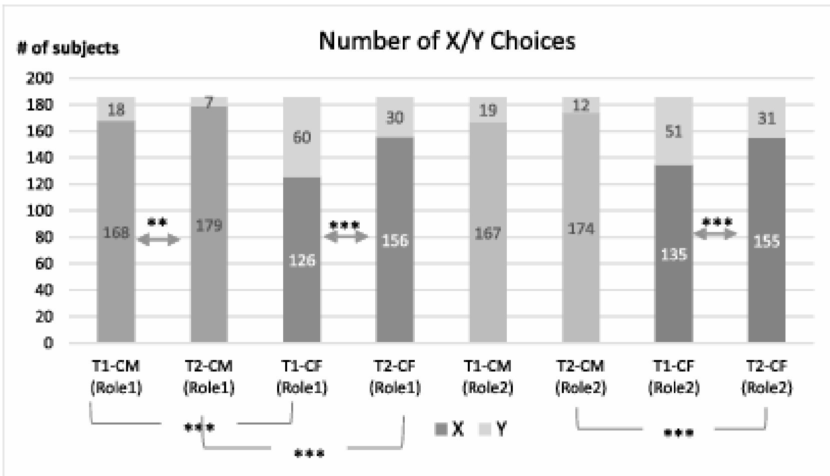
[22] The hypothesis that the four outcomes are randomly chosen with equal probabilities is rejected ($p = 0.01$).

[23] A multinomial logit regression over the four outcomes is no more informative than the descriptive statistics because of the skewed distribution of the choice data.

[24] The unique exception is the choice by the role 2 subjects who choose (Y, Y) less often in CF4 than in CF2 (Mann-Whitney U-test, $p = 0.1$).

Observation 1 *(Dictator decision) The behavior of the role 1 subjects in T0 is mostly consistent with self-interest and IEA. The behavior of the role 2 subjects in T0 is mostly consistent with self-interest and IQA.*[25]

5.2 Action choice



** and ***: significant difference at 5% and 1%, respectively, between the respective pair of distributions (χ^2-test). Shown in each column are the numbers of each action choice aggregated for k = 2, 4, and 6.

Figure 3: Action choices in T1 and T2

Figure 3 shows the frequency of each action (X and Y) by each subject role. As seen, there is a significant difference in the subjects' action choice in T1 and T2, and the difference is more prominent in CF: Going from CM-T1 to CM-T2, the choice of X increases by 6 percentage points for role 1 (90% $\rightarrow$ 96%) and by 4 percentage points for role 2 (90% $\rightarrow$ 94%). On the other hand, going from CF-T1 to CF-T2, the choice of X increases by 16 percentage points for role 1 (68% $\rightarrow$ 84%), and 10 percentage points for role 2 (73% $\rightarrow$ 83%). It also shows that both roles choose Y more often in CF than in CM. These observations are confirmed by the random effects logit regressions in Table 6, where the dependent variable equals one if a subject chooses Y. Models (4) and (6) in Table 6, which include inequality dummies k_4 and k_6, show that increasing inequality has different effects in CM and CF: While higher inequality overall has a positive impact on the choice of Y in CM, higher inequality has little to no impact in CF. In CM where the increasing inequality increases the choice of Y, this effect is independent of the subject role (model (5)). The observation on the action choice can be summarized as follows:

[25]Further analysis of role 2's choice in CF-T0 is given in Section 6.

Table 6: Random effect logit regressions of action choice: $y = \mathbf{1}_{\{a_i = Y\}}$

Model	(1)	(2)	(3)		(4)	(5)	(6)	(7)
	all	CM	CF		CM	CM	CF	CF
t2	-0.58	0.31	-0.71***	role	-0.07	0.60	0.38	-0.25
	(0.40)	(0.67)	(0.26)		(0.65)	(0.86)	(0.23)	(0.52)
cf	1.86***			k_4	0.92***	1.32**	0.21	-0.81*
	(0.30)				(0.15)	(0.60)	(0.40)	(0.48)
cf * t2	-0.23			k_6	1.11***	1.57***	-0.01	-0.01
	(0.37)				(0.29)	(0.58)	(0.40)	(0.50)
role		-0.06	0.38	role * k_4		-0.77		1.90***
		(0.53)	(0.23)			(1.20)		(0.65)
role * t2		-0.65	-0.44	role * k_6		-0.87		0.00
		(0.97)	(0.37)			(1.26)		(0.63)
1/round	6.55	22.00***	5.22	1/round	1.29**	1.32***	-0.12	-0.12
	-5.94	(7.34)	(6.62)		(0.51)	(0.46)	(0.41)	(0.44)
constant	-3.81***	-6.05***	-2.04***	Constant	-5.10***	-5.51***	-1.57***	-1.31***
	(0.73)	(0.85)	(0.60)		(1.13)	(0.94)	(0.33)	(0.36)
Log-likelihood	-513.09	-167.24	-342.10	Log-likelihood	-106.22	-105.93	-206.11	-201.55
#obs.	1,488	744	744	#obs.	372	372	372	372
#subjects	124	124	124	#subjects	124	124	124	124

Model (1) combines data from CM and CF whereas models (2)-(7) separate them. Independent variables: cf = 1 if CF, t2 = 1 if T2, role = 1 if role 1, $k_4 = 1$ if $k = 4$, and $k_6 = 1$ if $k = 6$. The variable 1/round equals the inverse of the round number within each task block, and is included given that all other independent variables are dummies. *, ** and ***: significant at 10%, 5% and 1%, respectively. Robust standard errors clustered by session in parentheses.

Observation 2 *(Action choice)*

1. *The choice of X is significantly less likely in CF than in CM.*
2. *In CF, T2 raises the choice of X compared with T1.*
3. *The subject role has no significant impact on the action choice in CM and CF.*
4. *Higher inequality increases the choice of Y in CM, while no such effect is observed in CF.*

Our central observation is Observation 2.2 on the comparison between T1 and T2, which cannot be explained by self-interest. The insignificance of the player role in CF in Observation 2.3 is in sharp contrast with our observation in T0, where the dominant choice is $A = (X, X)$ for role 1 and $D = (Y, Y)$ for role 2.[26]

5.3 Coordination

Table 7: Realization of action profiles

action profile	CM				CF			
	T0		T1	T2	T0		T1	T2
	Role 1	Role 2			Role 1	Role 2		
(X, X)	94	83	151	167	93	16	89	130
(Y, Y)	2	6	2	0	2	73	14	5
(X, Y)	0	1	17	12	1	0	37	26
(Y, X)	0	0	16	7	0	1	46	25
p-value (χ^2 test):								
T0=T1=T2	0.00	0.00			0.00	0.00		
T1=T2			0.07				0.00	
CM=CF		0.00	0.00	0.00				

The three lines in the bottom report the p-values of the χ^2 tests of the hypothesis that the distributions are the same for T0-T2 (first line), between T1 and T2 (second line), and between CM and CF (third line)

Table 7 describes the realized distribution of four action profiles in T0 through T2. It shows that the redistribution scheme induces efficient coordination particularly effectively in CF: Going from T1 to T2, the efficient coordination (X, X) increases by 9% percentage points (81% $\rightarrow$ 90%) in CM, but by 22 percentage points in CF (48% $\rightarrow$ 70%). Furthermore, the redistribution scheme also reduces coordination failures much more substantially in CF: Coordination failures (X, Y)

[26] In addition to some role 2 subjects who switch from D in T0 to X in CF-T1 or CF-T2, there are also some role 1 subjects who switch from $A = (X, X)$ in T0 to Y in CF-T1 or CF-T2. See more analysis on the behavior of individual subjects in Section 6, which also discusses the possible mechanism behind Observation 2.4.

and (Y, X) decrease by 8 percentage points in CM (18% → 10%), whereas they decrease by 18% percentage points in CF (45% → 27%). In fact, the difference in the distributions between T1 and T2 is strongly significant only in CF ($p = 0.00$) and only weakly so in CM ($p = 0.07$). The difference between CM and CF is strongly significant in T0, T1 and T2.

Table 8: Logit regressions of action profiles

Model	(1)	(2)	(3)	(4)	(5)	(6)
	All		CM		CF	
Dep. var.	$\mathbf{1}_{\{(X,X)\}}$	$\mathbf{1}_{\{(X,X) \text{ or } (Y,Y)\}}$	$\mathbf{1}_{\{(X,X)\}}$	$\mathbf{1}_{\{(X,X) \text{ or } (Y,Y)\}}$	$\mathbf{1}_{\{(X,X)\}}$	$\mathbf{1}_{\{(X,X) \text{ or } (Y,Y)\}}$
t2	0.81***	0.68***	1.09	1.06	0.85***	0.80***
	(0.23)	(0.22)	(0.79)	(0.78)	(0.33)	(0.21)
cf	-1.84***	-1.45***				
	(0.23)	(0.16)				
cf*t2	0.29	0.15				
	(0.38)	(0.35)				
1/round	-0.37	-0.40*	-0.64*	-0.45	0.02	-0.13
	(0.27)	(0.21)	(0.36)	(0.36)	(0.37)	(0.39)
k_4			-0.90***	-0.79***	0.00	0.23
			(0.15)	(0.13)	(0.48)	(0.55)
k_6			-0.98***	-0.85***	0.38	0.69**
			(0.24)	(0.28)	(0.45)	(0.34)
t2*k_4			0.26	0.10	0.43	0.06
			(0.96)	(0.94)	(0.79)	(0.79)
t2*k_6			-0.59	-0.68	0.34	-0.02
			(0.68)	(0.59)	(0.49)	(0.33)
constant	1.86***	1.84***	2.96***	2.78***	-0.24	-0.02
	(0.20)	(0.17)	(0.31)	(0.32)	(0.34)	(0.18)
#obs	744	744	372	372	372	372
Log likelihood	-374.40	-377.83	-136.58	-136.81	-233.66	-232.96

Models (1) and (2) combine data from CM and CF whereas models (3)-(6) separate them. See Table 6 for the definitions of the independent variables. *, ** and ***: significant at 10%, 5% and 1%, respectively. Robust standard errors clustered by session in parentheses.

These observations are again confirmed by logit regressions of action profiles in Table 8, where the dependent variable is either the efficient coordination ($\mathbf{1}_{\{a\,=\,(X,X)\}}$), or the total coordination ($\mathbf{1}_{\{a\,=\,(X,X) \text{ or } (Y,Y)\}}$).[27] As in the case of individual action choices, models (3)-(6) show that the impact of increasing inequality (*i.e.,* signs of the inequality dummies k_4 and k_6) is qualitatively different between CM and CF: higher inequality reduces coordination in CM, but either increases it or has no effect in CF. We summarize our findings as follows:

Observation 3 *(Coordination)*

[27] A multinomial logit regression is infeasible because of the heavily skewed frequency distribution of action profiles as indicated in Table 7.

1. *In both T1 and T2, higher inequality decreases coordination in CM, but has either positive or no effect on coordination in CF.*

2. *In both CM and CF, T2 increases efficient coordination (X, X) and reduces both inefficient coordination (Y, Y) and coordination failures.*

Observation 3.2 is our central finding, and in the case of CF, it corresponds directly to Observation 2.2 which finds the increased choice of action X in CF-T2 compared with CF-T1. The positive impact of high inequality on efficient coordination in CF (Observation 3.1) is something that is not evident from the analysis of the individual action choice (Observation 2.4).

5.4 Transfer

Table 9 shows the average transfer and the number of occurrences of positive transfers after each action profile. As seen, a dominant share of positive transfer is made by role 1: In total, role 1 makes positive transfers in 31.2% of all occasions in CM (58 times out of 186 occasions), and 43.5% of all occasions in CF (81 times out of 186 occasions). Role 1's average transfer is significantly positive in both CM and CF, implying that they are on average not self-interested. When aggregated over k, 94.8% and 84.0% of all positive transfers by role 1 are observed after the realization of (X, X) in CM and CF, respectively. Role 1's average transfer amount is significantly higher conditional on (X, X) than conditional on (X, Y) in both CM and CF ($p < 0.01$, t-test).

Table 15 in Appendix A.1 presents regressions of absolute and relative transfer as well as the likelihood of positive transfer by role 1 after his own choice of X given that positive transfer is observed almost exclusively in this combination as seen in Table 9. It shows that both the size of transfer and the likelihood of positive transfer are larger when role 2 chooses X in both CM and CF, and also larger in CF than in CM.[28] It shows that while absolute transfer increases with the inequality dummy k (models (4) and (7)), the inequality dummy k has no significant impact on the likelihood of positive transfer (models (6) and (9)).

Observation 4 *(Size and frequency of transfer)*

1. *The average transfer by role 1 is significantly positive.*

2. *Positive transfer by role 1 is more likely after the choice of X by role 2, and both absolute and relative transfer is larger in this case.*

3. *Positive transfer by role 1 is more likely in CF, and the size of transfer is larger in CF.*

[28]Relative transfer equals the absolute transfer amount divided by the payoff in the game: $\frac{t_1}{g_1}$. Figure 6 in Appendix A.1 also shows that CF dominates CM in terms of the cumulative distribution of relative transfer by role 1 ($p < 0.04$, Kolmogorov-Smirnov test).

Table 9: Average transfer ($\bar{t}_1, \bar{t}_2$) in T2 by game and action profile

	CM2				CM4				CM6			
	X		Y		X		Y		X		Y	
X	8.4	1.2	–	–	24.8	0.8	1.7	–	43.8	1.4	–	–
	$\frac{16}{59}$	$\frac{5}{59}$	$\frac{0}{3}$	$\frac{0}{3}$	$\frac{20}{56}$	$\frac{3}{56}$	$\frac{1}{3}$	$\frac{0}{3}$	$\frac{19}{52}$	$\frac{4}{52}$	$\frac{0}{6}$	$\frac{0}{6}$
Y	–	–	–	–	33.3	0.3	–	–	70.0	–	–	–
	$\frac{0}{0}$	$\frac{0}{0}$	$\frac{0}{0}$	$\frac{0}{0}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{0}{0}$	$\frac{0}{0}$	$\frac{1}{4}$	$\frac{0}{4}$	$\frac{0}{0}$	$\frac{0}{0}$
	CF2				CF4				CF6			
	X		Y		X		Y		X		Y	
X	10.9	0.6	6.7	–	27.9	1.0	–	–	42.4	2.5	0.3	–
	$\frac{17}{39}$	$\frac{2}{39}$	$\frac{2}{9}$	$\frac{0}{9}$	$\frac{24}{44}$	$\frac{3}{44}$	$\frac{0}{9}$	$\frac{0}{9}$	$\frac{27}{47}$	$\frac{5}{47}$	$\frac{1}{8}$	$\frac{0}{8}$
Y	4.6	1.8	–	–	15.8	0	–	–	44.0	0.2	–	–
	$\frac{4}{12}$	$\frac{2}{12}$	$\frac{0}{2}$	$\frac{0}{2}$	$\frac{4}{8}$	$\frac{0}{8}$	$\frac{0}{1}$	$\frac{0}{1}$	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{0}{2}$	$\frac{0}{2}$

The table lists for each action profile the average transfer amounts ($\bar{t}_1 : 1 \to 2$ and $\bar{t}_2 : 2 \to 1$) in line 1, and (#obs. of positive transfer)/(#obs. of the action profile) (by role 1 and role 2) in line 2.

4. *The inequality dummy k has a positive impact on absolute transfer, but has no impact on the likelihood of positive transfer.*

Observation 4.2 is the indication of reciprocity by role 1 in response to role 2's choice of the efficient NE action X. We will return to this point in Section 6.

5.5 Efficiency and equity

Our efficiency measure is the *total payoff,* which simply equals the sum of the two players' payoffs ($g_1 + g_2 = u_1 + u_2$). The introduction of redistribution increases the average total payoff by 4.7% in CM (490.65 in T1 ⇒ 513.92 in T2, $p = 0.28$ t-test), and by 12.8% in CF (301.51 in T1 ⇒ 340.00 in T2, $p = 0.03$ t-test). As seen in Figure 7 in Appendix A.1, the cumulative distribution of the total payoff in T2 (dashed-line) approximately first-order dominates that in T1 (solid-line) in both CM and CF, and the relationship is indeed significant in CF ($p = 0.07$, two-sided Kolmogorov-Smirnov test). In line with our earlier findings that the scheme facilitates efficient coordination (Observation 3.2) more strongly in CF, the redistribution scheme has a more substantial impact on efficiency in CF than in CM.

Further scrutiny of the subjects' payoffs reveals interesting facts. Figure 4 depicts the average final payoffs (u_i) in T1 and T2 for role 1 (dark) and role 2 (light). While redistribution raises role 2's payoff in both CM and CF ($p < 0.01$ in both

CM and CF, t-test and Mann-Whitney test), no such effect is observed for role 1 (by either test). Tobit regression analysis in Table 16 in Appendix A.1 confirms that redistribution has a significantly positive impact only on role 2's payoff. In summary, our finding suggests that role 1 transfers away any payoff gain from more efficient coordination achieved in T2.[29]

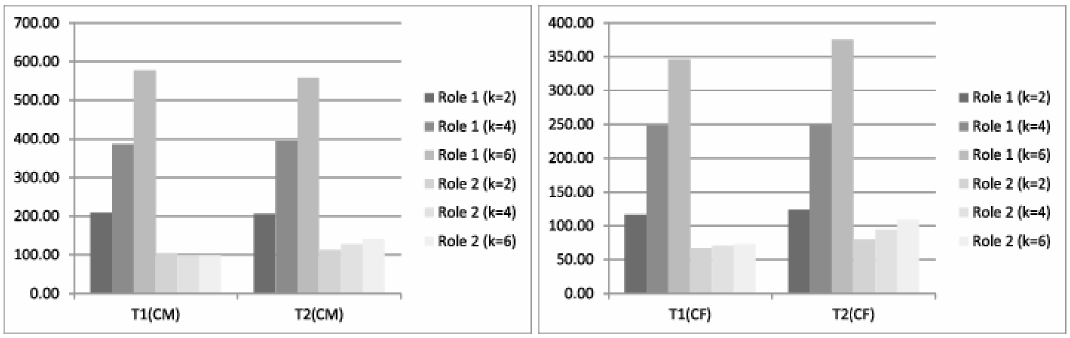


Figure 4: Final payoffs u_i in CM (left) and CF (right)

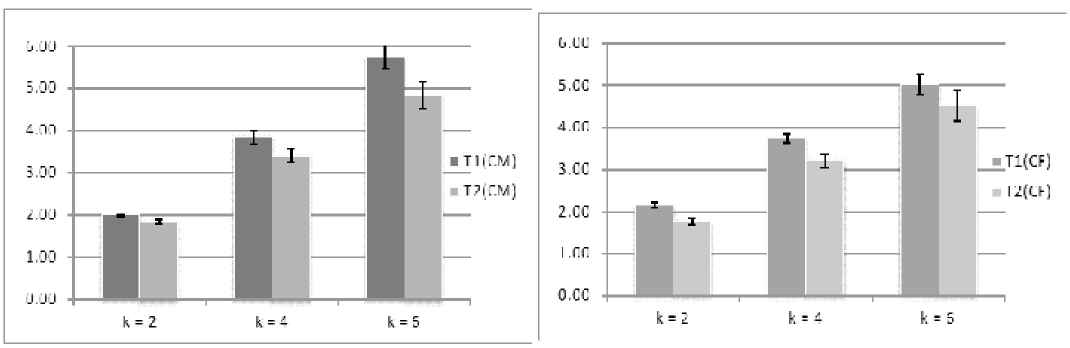


Figure 5: Final payoff ratios u_1/u_2 in CM (left) and CF (right)

Turning now to equity, we measure it by the ratio of the final payoffs:

$$\frac{u_1}{u_2} = \frac{\text{role 1's final payoff}}{\text{role 2's final payoff}}.$$

Figure 5 shows the average payoff ratio for T1 (dark) and T2 (light). For each k, we see that redistribution raises equity in both CM and CF. In fact, the null hypothesis of no difference between T1 and T2 is rejected for all k in CM ($p < 0.05$, t-test)

[29] One possible hypothesis behind this observation is that role 1 uses their payoff in T1 as the reference point and transfers away any additional gains in T2 to role 2. However, we find no support for this hypothesis as shown in Table 17 in Appendix A.1, which computes the average payoff of role 1 in T2 conditional on the action profile they experienced in T1.

and for $k = 2$ and $k = 4$ in CF ($p < 0.01$, t-test).[30] Strongly significant impact of redistribution on equity in both CM and CF is confirmed by the Tobit regressions of the payoff ratio in Table 18 in Appendix A.1.

Observation 5 *(Efficiency and equity)*

1. *Redistribution increases efficiency in CF, and the increased efficiency is almost entirely brought about by the increase in role 2's payoff.*

2. *Redistribution increases equity in both CM and CF.*

6 Discussion

How does the redistribution scheme induce efficient coordination and positive transfer? To answer this question, we first look at the aggregate data, and then investigate into the possible difference in motivations at the individual level.

As seen in Section 3, the inefficiency aversion (IEA) preferences imply no transfer in the redistribution stage, and hence fail to explain the positive transfer observed in our experiments. On the other hand, the inequality aversion (IQA) preferences always generate positive transfer by role 1, and the reciprocity preferences can also generate positive transfer. Regarding the comparison between IQA and reciprocity, IQA is supported in part by the observation that higher inequality (*i.e.,* a larger value of k) increases the size of the average transfer (Observation 4.4). On the other hand, however, the following findings support reciprocity but not IQA:

- The frequency of positive transfer by role 1 is higher after the choice of X by role 2 (Observation 4.2).

- Transfer is larger and more frequent in CF than in CM despite the larger payoff difference $g_1(X, X) - g_2(X, X)$ in CM than in CF (Observation 4.3).

- The frequency of positive transfer is unaffected by k (Observation 4.4).

The variation in the absolute and relative size of transfer across CF and CM, and across different values of the inequality dummy k, however, is not in line with our prediction based on reciprocity.[31] We summarize our observation as follows:

Observation 6 *The reciprocity preferences in (3) explain the presence of positive transfer by role 1 contingent on role 2's choice of X, and the larger frequency of positive transfer in CF than in CM.*

[30] The reduction in the payoff ratio in T2 is also significant by the Kolmogorov-Smirnov test ($p < 0.03$ for CM and $p < 0.01$ for CF).

[31] This is the consequence of our formulation of reciprocity in (3) which assumes that the degree of reciprocation is independent of the magnitude of the payoff increase from a to x.

We next look at the possible differences across subjects behind the increased choice of X in T2 compared with T1.

As observed earlier, some fraction of role 2 subjects choose $A = (X, X)$ in the dictator-decision task CF-T0. There are also subjects who switch from $D = (Y, Y)$ to A as inequality k increases. In other words, we can interpret these role 2 subjects as preferring efficiency to equity as the efficiency gap between A and D widens. We hence say that role 2 subjects are *inefficiency averse* (IEA) if, in CF-T0, they choose A for all k, or switch once from D to A as k increases. On the other hand, there are role 2 subjects who choose D for all levels of k, or switch from A to D once as k increases. These role 2 subjects are either self-interested or avoid an increase in inequality at A.[32] We call them *inequality averse* (IQA) type. Out of thirty role 2 subjects who made a choice in CF-T0, twenty are IQA, whereas six are IEA. Table 10 shows the difference in behavior of these types in T1 and T2. As seen, while type IEA chooses X most of the time and doesn't substantially change behavior from T1 to T2, type IQA chooses X less often overall, but increases the choice of X substantially in T2. As far as role 2 is concerned, hence, we can deduce that the increased choice of X in CF-T2 is by type IQA motivated by the reduced concern over inequality at (X, X) and/or the payoff loss at (X, X) (compared with (Y, Y)) in anticipation of the choice of X and positive transfer by role 1.

Table 10: Rate of action X in T1 and T2 by role 2's type in CF-T0

	CM-T1			CM-T2		
type	$k=2$	$k=4$	$k=6$	$k=2$	$k=4$	$k=6$
IQA	0.90	0.80	0.85	0.90	0.90	0.95
IEA	1.00	1.00	1.00	1.00	1.00	1.00

	CF-T1			CF-T2		
type	$k=2$	$k=4$	$k=6$	$k=2$	$k=4$	$k=6$
IQA	0.60	0.70	0.50	0.70	0.75	0.70
IEA	0.83	0.83	1.00	0.83	1.00	1.00

This inference is consistent with the subjects' responses to the post-experimental questionnaire: As for the choice of X in CF-T2, the most popular reason given by the role 2 subjects is "afraid of the low payoff resulting from my choice of Y and the other player's (role 1) choice of X," followed by "expecting the choice of X by role 1", "wanted to maximize the sum of payoffs", and "role 1 would reciprocate my choice of X with transfer." No role 2 subject describes altruistic motives such as wanting to be kind to role 1.[33] To see if role 2 is rational in their thinking, Table

[32]Role 2 subjects who choose D for every k can either be self-interested, or dislikes inequality even at $k = 2$. Based on CM-T0, only three subjects are classified as IQA. This implies that the classification here is specific to each class of games.

[33]Some of the role 1 subjects, on the other hand, thank role 2 for choosing X, implying the

19 in Appendix A.1 incorporates the average transfer $\bar{t}_1$ from role 1 to role 2 in T2 into their payoffs in CF. It shows that X is a dominant action for role 2 when $k = 6$, and (X, X) is a payoff dominant equilibrium when $k = 4$. Furthermore, if role 2 expects that role 1's action choice is given by its empirical frequency, X is his uniquely optimal action for every k.[34]

In T1 and T2, we also have subjects of both roles whose action choice either is constant for every k, or switches once as k goes up. In T1, we call a subject *type TX1* if his choice is X for every k, or switches once from Y to X as k increases. If his choice is Y for every k or switches once from X to Y as k increases, we call a subject *type TY1*. Define type TX2 and TY2 in T2 similarly.[35] TY is similar to type IQA in CF-T0 in that they respond to higher inequality with the choice of Y, and hence may be interpreted as inequality averse. Likewise, TX is similar to type IEA in CF-T0 and hence may be interpreted as inefficiency averse. As Facts 4 and 5 in Section 3.3 show, this interpretation is also consistent with the equilibrium behavior of inequality averse and inefficiency averse players in T1. It is important to note that behavior as described by types TX and TY reflects not only their preferences but also their beliefs over the strategic action choice by the other player. It is worth noting that the two types cover more than 90% of all cases in both T1 and T2.[36]

First, Table 11 shows the relationship between the IEA/IQA classification of role 2 subjects in CF-T0 and the TX1/TY1 classification of the same subjects in T1. As seen, type IEA in CF-T0 almost always becomes TX1 in both CM-T1 and CF-T1 so that their response to an increase in inequality is consistent across the two tasks. On the other hand, type IQA in CF-T0 is evenly split between TX1 and TY1 in CF-T1, and the predominant majority of them become TX1 in CM-T1. This also shows that the change in behavior of role 2 from T0 to T1 is brought about by the type who is concerned about inequality and/or the own payoff.

Table 12 shows the correlation of types in T1 and T2. In both CM and CF, nearly all of type TX1 become TX2 whereas approximately two-thirds of TY1 become TX2. In other words, those who respond to higher inequality with the choice of Y in T1 tend to change their behavior and respond to higher inequality with the choice of X in T2. Table 13 shows the likelihood of action X in T1 and T2 by each type in T1. As seen, type TX1 of either role chooses X most of the time in both T1 and T2 for each k. On the other hand, the change in the likelihood of X by type TY1

presence of reciprocity motives.

[34]For role 1, on the other hand, X cannot be a dominant action for any k. Again, however, if role 1 expects that role 2's action choice is given by its empirical frequency, X is the uniquely optimal action for role 1 for every k.

[35]The symbols TX and TY signify that the response moves **T**owards **X** and **T**owards **Y**, respectively, as k increases. These types are again specific to each class of games. The response of types TX1 and TX2 to k is hence similar to that of type IEA in T0 in the sense that efficiency is preferred over equity as k becomes larger, and that the response of types TY1 and TY2 is similar to that of type IQA in T0.

[36]The coverage is 90.3% in T1 (94.4% in CM-T1, 86.3% in CF-T1) and 94.4% in T2 (96.8% in CM-T2, 91.9% in CF-T2).

Table 11: Role 2 types in CF-T0 and T1

	CM-T1			CF-T1		
T0 types	TY1	TX1	other	TY1	TX1	other
IQA	3	15	2	9	10	1
IEA	0	6	0	0	5	1
other	0	3	1	0	4	0

is dramatic. While by definition they never choose X at $k = 6$ in T1, they choose X more than 60% of the time in T2.[37]

Table 13 also hints at the possible mechanism behind Observation 2.4 on the effect of inequality on the action choice. As seen, type TY1 in CM-T1 sharply decreases the choice of X as k goes up, whereas TX1 is mostly unresponsive to the change in k. If we associate TY1 with inequality-averse preferences as mentioned above, then we can interpret the negative impact of k on X in CM-T1 as resulting from inequality aversion of TY1. In CF-T1, on the other hand, TY1 chooses X less often than in CM-T1 overall, and TX1 increases their choice of X in response to an increase in k. Put differently, we can interpret that the positive effect of higher inequality on X in CF-T1 reflects the dominance of the inefficiency averse response by TX1 over the inequality averse response by TY1.

Observation 7 *Between T1 and T2, the increase in the choice of X by both roles 1 and 2 is mostly due to those subjects who respond to increased inequality with the choice of Y in T1. In T2, the choice of X by role 2 is motivated by the anticipation of role 1's choice of X* and *positive transfer.*

Table 12: Types in T1 and T2

	CM						CF					
	role 1			role 2			role 1			role 2		
	TY2	TX2	other	TY2	TX2	other	TY2	TX2	other	TY2	TX2	other
TY1	2	4	1	2	4	2	6	9	0	5	9	3
TX1	2	49	1	2	48	0	1	32	1	2	34	5
other	0	3	0	1	3	0	0	12	1	1	3	0

How is the difference in behavior in T1 translate to the difference in the transfer decisions in T2? As seen earlier, most positive transfers in T2 are by role 1 following the action profile (X, X). Table 20 in Appendix A.1 shows for both CM-T2 and CF-T2 the classification of role 1 subjects according to the number of times they experienced (X, X) and also the number of times they made positive transfer after

[37]On the other hand, the tendency of TY1 to increase the choice of Y for a larger k is unchanged even in T2.

Table 13: Rate of action X by T1 types

	CM-T1						CF-T1					
	role 1			role 2			role 1			role 2		
	$k=2$	$k=4$	$k=6$	$k=2$	$k=4$	$k=6$	$k=2$	$k=4$	$k=6$	$k=2$	$k=4$	$k=6$
TY1	0.86	0.57	0.00	0.75	0.50	0.00	0.47	0.20	0.00	0.53	0.35	0.00
TX1	0.94	0.98	1.00	0.98	1.00	1.00	0.79	0.91	1.00	0.78	1.00	1.00
	CM-T2						CF-T2					
	role 1			role 2			role 1			role 2		
	$k=2$	$k=4$	$k=6$	$k=2$	$k=4$	$k=6$	$k=2$	$k=4$	$k=6$	$k=2$	$k=4$	$k=6$
TY1	1.00	0.71	0.71	0.88	0.75	0.63	0.53	0.67	0.60	0.59	0.71	0.65
TX1	1.00	0.98	0.96	0.98	1.00	0.96	0.88	0.94	0.97	0.90	0.90	0.93

(X, X). As seen, in CF-T2, nearly half of role 1 subjects choose to make positive transfer every time they experience (X, X). In CM-T2, on the other hand, more than half of role 1 never make positive transfer after (X, X). Let the *reciprocation index* r be defined by

$$r = \frac{\#\text{positive transfers after } (X, X)}{\#\text{occurrences of } (X, X)}.$$

We call role 1 subjects *strongly reciprocating* (SR) if $r \geq \frac{2}{3}$, *weakly reciprocating* (WR) if $r \in \left[\frac{1}{3}, \frac{2}{3}\right)$, and *non-reciprocating* (NR) if $r < \frac{1}{3}$. Going from CF-T2 to CM-T2, we see overall downgrading of reciprocity types: Nearly half of type SR in CF-T2 become NR in CM-T2, while few type NR in CF-T2 become SR in CM-T2. Table 14 shows the relationship between the type classification in T1 (*i.e.,* TX1 and TY1 for role 1) and reciprocity types in T2. In CF-T2, the distribution of reciprocity types is almost identical between TY1 and TX1 with an even split between SR and NR. On the other hand, in CM-T2, nearly two-thirds of TX1 are NR, while it is much less likely that TY1 becomes NR. This shows that TY1 and TX1 are equally likely to reciprocate role 2's choice of X accompanied by payoff sacrifice, but that TX1 is more likely to ignore 2's choice of X if it is not accompanied by payoff sacrifice. In other words, TX1 and TY1 are different in the perception of kindness by role 2 that has led to the increase in the own payoff.

Fact 1 in Section 3.3 also implies that if role 1 reciprocates (X, X) with positive transfer at some k, then he does so at a larger k'. Table 21 in Appendix A.1 shows that this property holds mostly well in CF but less so in CM.[38]

Observation 8 *The degree of reciprocation is different between CM-T2 and CF-T2 even for the same subject. Those who respond to increased inequality by the choice*

[38] In CF, the likelihood of positive transfer is statistically different from 0.5 for each $k' > k$ when $k = 2$, 4, and 6 ($p < 0.01$ by the two-sided test of proportion). In CM, the likelihood is statistically different from 0.5 for each k' when $k = 2$ ($p < 0.05$) but not so when $k = 4$ or $k = 6$.

Table 14: Reciprocity types by type in T1

	CM			CF			
	SR	WR	NR	SR	WR	NR	other
TY1	5	0	2	5	1	5	4
TX1	12	9	31	16	3	15	0
other	1	2	0	7	2	4	0

"Other" refers to role 1 who didn't experience (X, X).

of X in T1 tends to not reciprocate role 2's choice of X not accompanied by payoff sacrifice.

The analysis so far is based on data from the six sessions which presented the payoff formula (1) in the instructions. There were five other sessions in which the instructions did not present the payoff formula.[39] Appendix A.2 reports some analysis that compares the results with and without the formula. Most notably, we observe that the inclusion of the formula had significantly positive impacts on the subjects' action choice both in T1 and T2. In particular, inclusion of the formula decreases the frequency of action Y and increases the frequency of efficient coordination (X, X). These results suggest that inclusion of the formula raises the awareness of the externalities involved in decision making in the inequality games. We believe that such awareness of externalities is key to the inducement of social preferences including reciprocity.[40]

7 Conclusion

We study the working of the redistribution scheme, which allows ex post voluntary transfer of payoffs, in a class of 2×2 coordination games with an efficient NE (X, X) that favors player 1 over player 2, and an equitable but inefficient NE (Y, Y). We find that the redistribution scheme induces positive transfer from player 1 to player 2, and also increases efficient coordination (X, X). Through the analysis of the size and frequency of transfer, we conclude that the positive transfer is likely a reciprocative response by player 1 to player 2's choice of action X. Importantly, we find that the scheme is more effective in the games in which player 2's payoff at (X, X) is lower than that at (Y, Y), and interpret this as suggesting the stronger

[39] These sessions were identical to the main sessions otherwise. The five sessions had 20, 20, 20, 24, and 22 subjects with the total of 106 subjects.

[40] As mentioned in Footnote 19, we had four pilot sessions with rotated task orders. These sessions also presented no payoff formula in the instructions. Analysis combining data from all sessions without the payoff formula shows that the effect of rotation on the likelihood of coordination is insignificant, but tends to be negative. The effect of T2 on coordination is in line with our main analysis.

reciprocity shown by player 1 when 2's choice of X entails self sacrifice. Analysis of subjects' individual behavior suggests that player 2 chooses X in anticipation of the choice of X and positive transfer by player 1. Furthermore, the increase in the choice of X under the redistribution scheme is attributed to the type who, in the absence of ex post redistribution, responds to higher inequality with the choice of Y.

One interesting extension of the present work involves elicitation of beliefs before the play of the games. It would be interesting to find out beliefs about the other player's action choice, and the amount of transfer they expect from the other player after the realization of each action profile. Such study would provide us with more insight into the working of the redistribution scheme. Although we have studied the redistribution scheme in the presence of inequality between the players, it is important to check its validity in other classes of games. In the BOS game, for example, we would expect that the redistribution scheme as proposed here is valid if the sum of payoffs at one NE is substantially higher than that at the other NE. If, on the other hand, both NE are equally efficient, then some modification to the scheme would be required. For example, restricting transfer to one way may be one possible solution in such a case. Examining the validity of the scheme under various payoff specifications is a topic of future research.

References

N. Anbarcı, N. Feltovich, and M. Y. Gürdal. Payoff inequity reduces the effectiveness of correlated-equilibrium recommendations. *European Economic Review*, 108(C): 172–190, 2018.

J. Andreoni, W. Harbaugh, and L. Vesterlund. The carrot or the stick: Rewards, punishments, and cooperation. *American Economic Review*, 93(3):893–902, June 2003.

L. Belafoutas, M. G. Kocher, L. Putterman, and M. Sutter. Equality, equity and incentives: An experiment. *European Economic Review*, 60:32–51, 2013.

J. Bone, M. Drouvelis, and I. Ray. Coordination in 2×2 games by following recommendations from correlated equilibria. Working Paper, 2013.

E. Buckley and R. Croson. The poor give more: Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics*, 90(5):935–955, 2006.

G. P. Cachon and C. F. Camerer. Loss-avoidance and forward induction in experimental coordination games. *Quarterly Journal of Economics*, 111(1):165–194, 1996.

T. N. Cason and T. Sharma. Recommended play and correlated equilibria: an experimental study. *Economic Theory*, 33(1):11–27, Oct 2007.

G. Charness and M. Dufwenberg. Promises and partnership. *Econometrica*, 74(6): 1579–1601, 2006.

R. Cooper, D. V. DeJong, R. Forsythe, and T. W. Ross. Communication in coordination games. *Quarterly Journal of Economics*, 107(2):739–771, 1992.

R. Cooper, D. V. DeJong, R. Forsythe, and T. W. Ross. Forward induction in the battle-of-the-sexes games. *American Economic Review*, 83(5):1303–1316, 1993.

R. W. Cooper, D. V. DeJong, R. Forsythe, and T. W. Ross. Selection criteria in coordination games: Some experimental results. *American Economic Review*, 80 (1):218–233, 1990.

V. P. Crawford, U. Gneezy, and Y. Rottenstreich. The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, 98(4):1443–1458, 2008.

S. Dekel, S. Fischer, and R. Zultan. Potential pareto public goods. *Journal of Public Economics*, 146:87–96, 2017.

J. Duffy and N. Feltovich. Correlated equilibria, good and bad: An experimental study. *International Economic Review*, 51(3):701–721, 2010.

M. Dufwenberg and G. Kirchsteiger. A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268–298, 2004.

N. Erkal, L. Gangadharan, and N. Nikiforakis. Relative earnings and giving in a real-effort experiment. *American Economic Review*, 101:3330–3348, 2011.

P. Evdokimov and A. Rustichini. Forward induction: Thinking and behavior. *Journal of Economic Behavior & Organization*, 128:195–208, 2016.

E. Fehr and S. Gächter. Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4):980–994, September 2000.

E. Fehr and B. Rockenbach. Detrimental effects of sanctions on human altruism. *Nature*, 422(6928):137–40, Mar 13 2003.

U. Fischbacher. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178, 2007.

L. Gangadharan, N. Nikiforakis, and M. C. Villeval. Normative conflict and the limits of self-governance in heterogeneous populations. *European Economic Review*, 100:143–156, 2017.

J. K. Goeree and C. A. Holt. An experimental study of costly coordination. *Games and Economic Behavior*, 51:349–364, 2005.

A. Hofmeyr, J. Burns, and M. Visser. Income inequality, reciprocity and public good provision: An experimental analysis. *South African Journal of Economics*, 75:508–520, 2007.

D. Houser, E. Xiao, K. McCabe, and V. Smith. When punishment fails: Research on sanctions, intentions and non-cooperation. *Games and Economic Behavior*, 62 (2):509–532, 2008.

S. Huck and W. Müller. Burning money and (pseudo) first-mover advantages: an experimental study on forward induction. *Games and Economic Behavior*, 51(1): 109–127, 2005.

D. Masclet, C. Noussair, S. Tucker, and M.-C. Villeval. Monetary and nonmonetary punishment in the voluntary contributions mechanism. *American Economic Review*, 93(1):366–380, March 2003.

F. Ohtake, Y. Kinari, N. Mizutani, and T. Mori. Income, giving, and egalitarianism: A real-effort experiment in Japan. *Journal of Behavioral Economics and Finance*, 6:81–84, 2013. (in Japanese).

R. J. Oxoby and J. Spraggon. A clear and present minority: Heterogeneity in the source of endowments and the provision of public goods. *Economic Inquiry*, 51 (4):2071–2082, 2013.

M. Rabin. Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):1281–1302, 1993.

P. G. Straub. Risk dominance and coordination failures in static games. *Quarterly Review of Economics and Finance*, 35(4):339–363, 1995.

N. Uler. Public goods provision, inequality and taxes. *Experimental Economics*, 14: 287–306, 2011.

J. B. Van Huyck, R. C. Battalio, and R. O. Beil. Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*, 80(1):234–248, March 1990.

Appendix

A.1 Figures and Tables

Table 15: Determinants of the size and likelihood of transfer

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		All			CM			CF	
	absolute	relative	likelihood	absolute	relative	likelihood	absolute	relative	likelihood
2's Y	-149.14***	-0.28***	-2.16***						
	(7.71)	(0.01)	(0.27)						
cf	23.54**	0.08***	1.06***						
	(10.20)	(0.02)	(0.27)						
2's Y * cf	28.21	0.132*	-0.23						
	(31.12)	(0.08)	(0.67)						
k_4				40.83***	0.04**	0.31	32.53**	0.00	0.33
				(12.88)	(0.02)	(0.29)	(15.68)	(0.03)	(0.42)
k_6				81.93***	0.07***	0.38	51.14***	0.00	0.56
				(13.29)	(0.02)	(0.26)	(13.96)	(0.03)	(0.49)
1/round	23.64	0.0469*	0.81*	46.44**	0.10***	1.48***	23.92	-0.01	0.16
	(18.79)	(0.03)	(0.48)	(19.81)	(0.03)	(0.55)	(22.00)	(0.05)	(0.79)
constant	-48.33***	-0.10***	-1.22***	-134.15***	-0.20***	-2.00***	-54.43**	-0.03	-0.62
	(18.26)	(0.04)	(0.37)	(13.34)	(0.02)	(0.38)	(24.99)	(0.05)	(0.70)
#obs.	335	335	335	179	179	179	156	156	156
Log-likelihood	-808.31	-17.03	-157.29	-379.61	-23.88	-87.71	-432.53	-11.46	-88.52

Models (1), (2), (4), (5), (7) and (8) are the Tobit regressions of the relative and absolute transfer amounts, whereas (3), (6) and (9) are the probit regressions of the likelihood $\mathbf{1}_{\{t_1>0\}}$ of positive transfer. The variable "2's Y"= 1 if role 2's action is Y. *, ** and ***: significant at 10%, 5% and 1%, respectively. Robust standard errors clustered by session in parentheses.

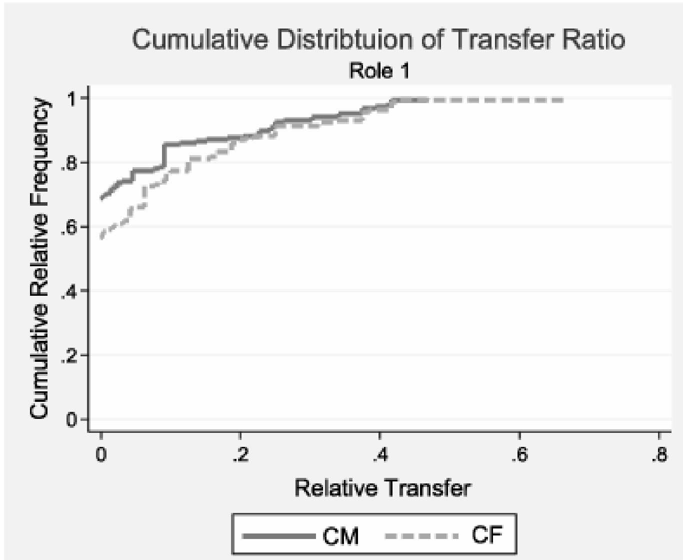


Figure 6: Cumulative distributions of relative transfer by role 1 subjects

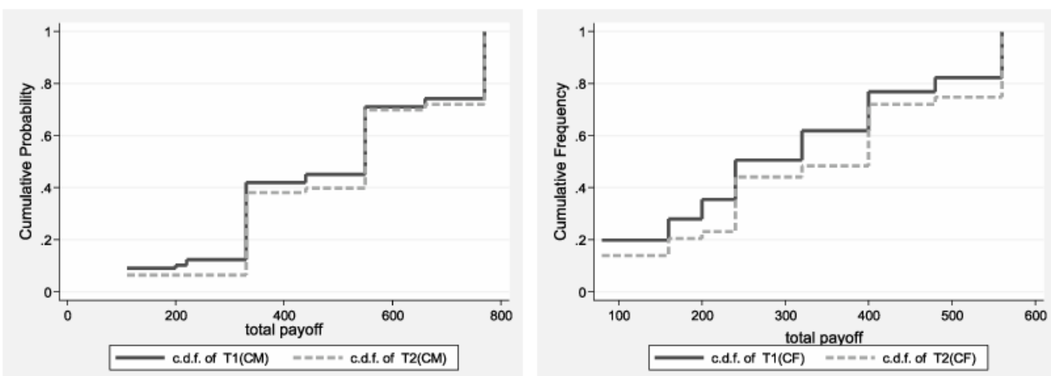


Figure 7: Cumulative distributions of total payoffs in T1 and T2: CM (left) and CF (right)

Table 16: Mixed effect Tobit regressions of final payoffs

	Role 1		Role 2	
	CM	CF	CM	CF
Model	(1)	(2)	(3)	(4)
t2	-3.12	7.691	10.08***	12.56***
	(6.48)	(5.33)	(2.48)	(1.78)
k_4	174.40***	132.10***	-5.61***	3.07
	(11.45)	(7.22)	(1.18)	(4.10)
k_6	367.60***	230.10***	-5.63***	5.86***
	(18.49)	(10.53)	(1.76)	(1.60)
t2 * k_4	12.18	-6.034	19.68***	11.44
	(10.65)	(16.43)	(4.99)	(7.42)
t2 * k_6	-15.65	21.00	32.65***	23.03***
	(23.98)	(18.25)	(6.54)	(5.39)
1/round	12.01	0.713	-0.13	1.75
	(19.92)	(11.45)	(4.60)	(3.72)
constant	204.80***	115.90***	103.90***	66.93***
	(6.11)	(6.20)	(2.71)	(0.80)
# of obs.	372	372	372	372
Log likelihood	-2333.37	-2283.6033	-1944.21	-1839.3891

*, ** and ***: significant at 10%, 5% and 1%, respectively. Standard errors clustered by session in parentheses.

Table 17: Role 1's payoff in T2 conditional on the action profile in T1

T1	CF2	CF4	CF6	CM2	CM4	CM6
(X,X)	132.429	231.393	384.857	206.855	405.438	560.208
	(40.547)	(106.812)	(136.046)	(34.280)	(85.694)	(166.431)
	28	28	35	55	48	48
	0.001	0.002	0.002	0.0106	0.008	0.001
(X,Y)	134.059	266.375	472.727	166.667	345.143	565.714
	(34.965)	(92.751)	(13.484)	(92.376)	(137.914)	(224.117)
	17	8	11	3	7	7
	0	0.0004	0	0.1835	0.0016	0.001
(Y,X)	99.800	275.286	322.444	214.500	366.667	548.333
	(31.134)	(65.198)	(193.891)	(9.713)	(78.655)	(153.677)
	15	21	9	4	6	6
	0.9805	0.2954	0.1696	–	–	–
(Y,Y)	100.000	220.000	337.143	–	438.000	460.000
	(0.000)	(88.318)	(142.912)	–	–	–
	2	5	7	–	1	1
	–	–	–	–	–	–

For each action profile in T1, the table lists the average payoff in T2 (line 1), standard deviations (line 2), the number of observations (line 3), and *p*-value of the hypothesis: "payoff in T1= payoff in T2" by t-test (line 4). "–" implies insufficient observations.

Table 18: Tobit regressions of the payoff ratio

VARIABLES	(1) CM	(2) CM	(3) CM	(4) CF	(5) CF	(6) CF
t2	-0.51***	-0.51***	-0.17***	-0.47***	-0.47***	-0.37***
	(0.15)	(0.16)	(0.05)	(0.15)	(0.11)	(0.06)
k4		1.64***	1.79***		1.51***	1.59***
		(0.18)	(0.16)		(0.12)	(0.18)
k6		3.36***	3.73***		2.80***	2.85***
		(0.24)	(0.30)		(0.13)	(0.05)
1/round	0.88	0.29	0.26	-0.87	-0.01	-0.05
	(0.77)	(0.50)	(0.40)	(0.85)	(0.11)	(0.15)
t2 * k4			-0.28			-0.17
			(0.20)			(0.17)
t2 * k6			-0.73**			-0.11
			(0.35)			(0.27)
constant	3.51***	2.09***	1.93***	3.98***	2.21***	2.18***
	(0.09)	(0.28)	(0.12)	(0.34)	(0.06)	(0.08)
Log likelihood	-794.54	-681.63	-679.59	-768.93	-686.67	-686.56
#obs	372	372	372	372	372	372
#subject pairs	62	62	62	62	62	62

*, ** and ***: significant at 10%, 5% and 1%, respectively. Robust standard errors clustered by session in parentheses.

Table 19: Payoffs incorporating the average transfer from role 1: $(g_1 - \bar{t}_1, g_2 + \bar{t}_1)$

CF2					CF4					CF6				
	X (51)		Y (11)			X (52)		Y (10)			X (52)		Y (10)	
X(48)	149.1,	90.9	53.3,	26.7	X(53)	292.1,	107.9	60,	20	X(55)	437.6,	122.4	59.7,	20.3
Y(14)	95.4,	64.6	100,	100	Y(9)	244.3,	75.8	100,	100	Y(7)	376,	104	100,	100

#observations in parentheses.

A.2 Effect of the Payoff Formula in the Instructions

This section examines the effects of including the payoff formula (1) in the instructions. There are a total of 106 subjects who participated in five sessions without the payoff formula but with the same task sequence as in the main experiments. Tables 22 and 23 describe the frequency of action Y by each role and the frequency of each action profile, respectively, in T1 and T2 with and without the payoff formula. We observe that role 1 chooses Y less often in every game with the formula, and that role 2 does so in four out of six games (CF2 and CM6). The effect is stronger in T2. In the case of action profiles, the efficient coordination profile (X, X) increases with the formula in every game, whereas the inefficient coordination profile decreases or does not change with the formula in every game. Again, these effects are generally stronger in T2. As seen in logit regressions reported in Tables 24 and 25, many of these changes are significant. In terms of transfer, however, the inclusion of the payoff formula has no positive impact on the average transfer by role 1 as seen in

Table 20: Positive transfer by role 1 following (X, X)

		$\#\{t_1 > 0\}$							
		CM-T2				CF-T2			
		0	1	2	3	0	1	2	3
$\#(X, X)$	0	0	-	-	-	4	-	-	-
	1	2	0	-	-	4	5	-	-
	2	5	5	5	-	11	4	11	-
	3	26	6	5	8	9	2	1	11

Table 21: Inequality and positive transfer

		Positive transfer after (X,X) #obs	%positive transfer after (X,X) $k = 2$	$k = 4$	$k = 6$
CM-T2	$k = 2$	16		79	92
	$k = 4$	20	61		69
	$k = 6$	19	67	69	
CF-T2	$k = 2$	17		93	93
	$k = 4$	24	81		100
	$k = 6$	27	87	86	

The second column shows the number of role 1 subjects who made positive transfer after (X, X) for each k, and the three right columns show what percentage of them made positive transfer after (X, X) for different levels of k.

the logit regressions reported in Table 26.

On the other hand, the redistribution scheme increases the choice of X even without the formula: Going from T1 to T2, the rate of X increases by 6.8 percentage points in CM (83% $\rightarrow$ 89.3% for role 1 and 83% $\rightarrow$ 90.3% for role 2), and by 13.2 percentage points in CF (57.3% $\rightarrow$ 67.3% for role 1 and 65.3% $\rightarrow$ 81.7% for role 2). However, the increase is smaller than the corresponding number with the formula reported in Section 5.5.2.

Table 22: Frequencies of Y with and without formula

	role 1						role 2					
T1	CF2	CF4	CF6	CM2	CM4	CM6	CF2	CF4	CF6	CM2	CM4	CM6
without	0.34	0.51	0.43	0.19	0.15	0.17	0.38	0.28	0.38	0.09	0.21	0.21
53	(0.07)	(0.07)	(0.07)	(0.54)	(0.05)	(0.05)	(0.07)	(0.06)	(0.07)	(0.04)	(0.06)	(0.06)
with	0.27	0.42	0.27	0.06	0.11	0.11	0.31	0.21	0.31	0.05	0.13	0.13
62	(0.06)	(0.06)	(0.06)	(0.03)	(0.04)	(0.04)	(0.06)	(0.05)	(0.06)	(0.03)	(0.04)	(0.04)
T2												
without	0.38	0.26	0.34	0.06	0.13	0.13	0.17	0.17	0.21	0.09	0.09	0.11
53	(0.07)	(0.06)	(0.07)	(0.03)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.04)	(0.04)	(0.04)
with	0.23	0.15	0.11	0.00	0.05	0.06	0.18	0.16	0.16	0.05	0.05	0.10
62	(0.05)	(0.04)	(0.04)	(0.00)	(0.03)	(0.03)	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)	(0.04)

Standard errors in parentheses

Table 23: Action profiles with and without formula

		CF2		CF4		CF6	
		without	with	without	with	without	with
T1	(X, X)	0.42	0.45	0.32	0.45	0.38	0.53
	(X, Y)	0.25	0.27	0.17	0.13	0.19	0.19
	(Y, X)	0.21	0.24	0.40	0.34	0.25	0.16
	(Y, Y)	0.13	0.03	0.11	0.08	0.19	0.11
	Fisher's test	0.365		0.542		0.307	
T2	(X, X)	0.53	0.63	0.60	0.71	0.55	0.76
	(X, Y)	0.09	0.15	0.13	0.15	0.11	0.13
	(Y, X)	0.30	0.19	0.23	0.13	0.25	0.08
	(Y, Y)	0.08	0.03	0.04	0.02	0.09	0.03
	Fisher's test	0.32		0.459		0.033	
		CM2		CM4		CM6	
T1	(X, X)	0.77	0.89	0.64	0.77	0.64	0.77
	(X, Y)	0.04	0.05	0.21	0.11	0.19	0.11
	(Y, X)	0.13	0.06	0.15	0.10	0.15	0.10
	(Y, Y)	0.06	0.00	0.00	0.02	0.02	0.02
	Fisher's test	0.145		0.274		0.472	
T2	(X, X)	0.87	0.95	0.81	0.90	0.79	0.84
	(X, Y)	0.08	0.05	0.06	0.05	0.08	0.10
	(Y, X)	0.04	0.00	0.09	0.05	0.09	0.06
	(Y, Y)	0.02	0.00	0.04	0.00	0.04	0.00
	Fisher's test	0.254		0.318		0.409	

Table 24: Logit regression of action choice Y with and without formula

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	CF role1	CF role1	CM role1	CM role1	CF role2	CF role2	CM role2	CM role2
formula	-1.141**	-0.506	-1.386***	-0.855*	-0.825**	-0.445	-1.038	-0.749
	(0.50)	(0.52)	(0.46)	(0.45)	(0.38)	(0.36)	(0.69)	(0.71)
t2	-0.057	-0.042	-0.0808**	-0.0699**	-0.128***	-0.118***	-0.0630**	-0.0564*
	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
1/period	0.842***	0.893***	1.185**	1.227**	-1.062***	-1.082***	-0.376	-0.347
	(0.30)	(0.32)	(0.60)	(0.57)	(0.35)	(0.38)	(0.70)	(0.70)
t2 * formula		-1.405***		-1.371***		-0.804***		-0.625
		(0.31)		(0.39)		(0.16)		(0.61)
Constant	-0.937**	-1.035**	-3.442***	-3.546***	-0.552	-0.585	-2.522***	-2.566***
	(0.46)	(0.47)	(0.63)	(0.64)	(0.36)	(0.37)	(0.61)	(0.64)
Log likelihood	-351.95	-342.23	-178.54	-175.15	-332.86	-329.45	-198.85	-197.88
#obs	690	690	690	690	690	690	690	690
#subjects	115	115	115	115	115	115	115	115

The variable formula = 1 for sessions with the formula. *, ** and ***: significant at 10%, 5% and 1%, respectively. Robust standard errors clustered by session in parentheses.

Table 25: Logit regression of action profile with and without formula

VARIABLES	(1) CF yy	(2) CF yy	(3) CM yy	(4) CM yy	(5) CF xx/yy	(6) CF xx/yy	(7) CM xx/yy	(8) CM xx/yy
formula	-1.428***	-0.950*	-1.727**	-0.99	0.417**	0.02	0.918**	0.596
	(0.54)	(0.56)	(0.83)	(0.86)	(0.17)	(0.18)	(0.37)	(0.37)
t2	-0.09	-0.08	0.00	0.01	0.04	0.03	0.0766***	0.0695***
	(0.07)	(0.07)	(0.05)	(0.05)	(0.03)	(0.03)	(0.01)	(0.01)
1/period	-1.563***	-1.637***	1.405**	1.416**	-0.431**	-0.442**	-0.113	-0.129
	(0.42)	(0.43)	(0.70)	(0.65)	(0.20)	(0.20)	(0.38)	(0.36)
t2 * formula		-1.303***		omitted		0.772***		0.696**
		(0.30)				(0.21)		(0.27)
Constant	-2.006***	-2.045***	-4.784***	-4.832***	0.340	0.383*	1.395***	1.432***
	(0.57)	(0.58)	(0.55)	(0.56)	(0.22)	(0.22)	(0.51)	(0.52)
Log likelihood	-166.33	-163.64	-51.63	-50.20	-456.03	-450.20	-298.47	-296.31
#obs	690	690	690	504	690	690	690	690
#subjects	115	115	115	115	115	115	115	115

*, ** and ***: significant at 10%, 5% and 1%, respectively. Standard errors clustered by session in parentheses.

Table 26: Transfer with and without formula

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	All data												conditional on xx			
					conditional on xx				role 1	role 2	role 1	role 2	role 1	role 2	role 1	role 2
VARIABLES	role 1	role 1	role 2	role 2	role 1 xx	role 1 xx	role 2 xx	role 2 xx	CF	CF	CM	CM	CF	CF	CM	CM
formula	13.43	15.33	-11.97***	-13.04***	14.67	14.78	-10.47**	-12.81***	9.783	-15.51*	11.85	-8.047***	5.154	-12.42	13.43	-9.235***
	(9.20)	(11.47)	(4.57)	(4.74)	(9.57)	(13.06)	(4.35)	(4.60)	(9.33)	(8.84)	(14.26)	(2.63)	(9.29)	(10.24)	(15.00)	(2.49)
cf	11.56**	13.50***	-0.51	-1.31	22.38***	22.53***	-2.495	-5.095***								
	(4.55)	(3.49)	(2.26)	(3.48)	(6.17)	(4.54)	(1.61)	(1.29)								
1/period	24.27*	24.24*	5.242**	5.179**	23.66*	23.66*	5.697	5.202	12.73	11.29***	54.81***	2.729	7.426	15.15**	49.87***	-0.537
	(13.37)	(13.38)	(2.48)	(2.45)	(12.39)	(12.49)	(4.06)	(4.02)	(12.14)	(3.94)	(11.65)	(4.46)	(11.88)	(7.02)	(14.68)	(4.83)
cf * formula		-3.482		2.175		-0.248		6.348**								
		(8.62)		(3.53)		(10.39)		(2.88)								
k_4									30.81***	-7.610*	34.79***	-3.178	23.73***	-3.049	35.92***	-3.596
									(7.74)	(3.91)	(8.43)	(2.43)	(7.20)	(4.99)	(9.93)	(3.37)
k_6									46.27***	0.566	71.80***	-2.943	42.07***	1.995	77.56***	-1.595
									(9.34)	(4.39)	(13.88)	(3.15)	(9.72)	(7.27)	(12.11)	(2.89)
role																
role * formula																
Constant	-66.60***	-67.66***	-33.00***	-32.56***	-60.93***	-60.99***	-32.64***	-31.36***	-58.04***	-33.24***	-141.8***	-33.88***	-35.77***	-35.02***	-135.4***	-29.80***
	(10.46)	(10.00)	(7.35)	(7.19)	(10.25)	(10.08)	(7.43)	(6.95)	(11.84)	(9.04)	(16.97)	(7.64)	(12.69)	(11.19)	(17.09)	(6.32)
Log likelihood	-1597.71	-1597.67	-495.73	-495.61	-1318.36	-1318.36	-393.61	-392.55	-890.71	-266.30	-699.12	-256.71	-661.38	-177.50	-633.06	-239.28
#obs	690	690	690	690	517	517	517	517	345	345	345	345	219	219	298	298
#subject pairs	115	115	115	115	113	113	112	112	115	115	115	115	99	101	113	112

*, ** and ***: significant at 10%, 5% and 1%, respectively. Standard errors clustered by session in parentheses.

A.3 Proofs

Proof of Proposition 1. The utility function U_i is concave in the own transfer t_i so that the first-order condition fully characterizes the solution to the maximization problem. In particular, the solution is either at a corner ($t_i = 0$ or $t_i = g_i$) or in the interior ($t_i = \gamma_i(x) - g_j(x) + t_j$). When $\gamma_1(x) + \gamma_2(x) \neq g_1(x) + g_2(x)$, we cannot have both t_1 and t_2 as interior solutions. The first-order condition for t_i against $t_j = 0$ or $t_j = g_j$ then yields the relationship between (γ_1, γ_2) and t_i. Given (4), neither player transfers his entire payoff at (Y, X) if $\gamma_1(Y, X) + \gamma_2(Y, X) = \nu_1 + 0 < b + c_1$, and at (X, X) if $\gamma_1(X, X) + \gamma_2(X, X) = \mu_1 + \mu_2 < 2b + c_1 + c_2$ in CM, and $\gamma_1(X, X) + \gamma_2(X, X) = \nu_1 + 0 < 2b + c_1 + c_2$ in CF. When these conditions hold, only player 1 makes a transfer at (Y, X) if $\gamma_1(Y, X) = \nu_1 > g_2(Y, X) = b$. In this case, 2's payoff at (Y, X) equals $U_2(Y, X) = \nu_1$, and hence $x_2 = X$ is 2's best response against $x_1 = Y$ if $\nu_1 > a$. In this case, $U_1(X, X) > U_1(Y, X)$ since $g_1(X, X) > g_1(Y, X)$ and since $\mu_1 \leq \nu_1$. Furthermore, only player 1 makes a transfer at (X, X) if $\gamma_1(X, X) = \mu_1 > b + c_2$ in CM, and if $\gamma_1(X, X) = \nu_1 > b + c_2$ in CF. ■

Equilibrium under distributive social preferences Let e_i^{T0} denote player i's optimal choice in the dictator task T0, and E^{T1} and E^{T2} denote the set of (pure) NE and SPE action profiles in the inequality game G in tasks T1 and T2, respectively.

1) Inefficiency aversion

In T0, the optimal action for player 1 is $e_1^{T0} = (X, X)$ regardless of κ_1, and for player 2,

$$e_2^{T0} = \begin{cases} (X, X) & \text{if } \kappa_2 > \frac{a-b-c_2}{2b+c_1+c_2-2a}, \\ (Y, Y) & \text{if } \kappa_2 < \frac{a-b-c_2}{2b+c_1+c_2-2a}. \end{cases}$$

In T1,

$$E^{T1} = \begin{cases} \{(X, X)\} & \text{if } b + c_1 > 2a \text{ and } \kappa_2 > \frac{a-b}{b+c_1-2a},^{41} \\ \{(X, X), (Y, Y)\} & \text{if } b + c_1 \leq 2a, \text{ or if } b + c_1 > 2a \text{ and } \kappa_2 < \frac{a-b}{b+c_1-2a}. \end{cases} \tag{7}$$

In T2, the transfer equals zero $t \equiv 0$ and the set of SPE action profiles is as given in (7): $E^{T2} = E^{T1}$.

2) Inequality aversion

In T0,

$$e_1^{T0} = \begin{cases} (X, X) & \text{if } \lambda_1 < \frac{b+c_1-a}{c_1-c_2}, \\ (Y, Y) & \text{if } \lambda_1 > \frac{b+c_1-a}{c_1-c_2}, \end{cases}$$

[41] $b + c_1 > 2a$ holds in all but one (CF1) of our parameter specifications. See Table 3.

and

$$e_2^{T0} = \begin{cases} (X,X) & \text{if } \lambda_2 < \frac{b+c_2-a}{c_1-c_2}, \\ (Y,Y) & \text{if } \lambda_2 > \frac{b+c_2-a}{c_1-c_2}, \end{cases}$$

In T1,

$$E^{T1} = \begin{cases} \{(Y,Y)\} & \text{if } \lambda_1 > \frac{b}{b-c_2} \text{ or } \lambda_2 > \frac{b}{c_1-b}, \\ \{(X,X),(Y,Y)\} & \text{if } \lambda_1 \le \frac{b}{b-c_2} \text{ and } \lambda_2 \le \frac{b}{c_1-b}. \end{cases}$$

In T2, if $\lambda_1 < \frac{1}{2}$, then no transfer takes place in SPE, and the SPE action profile in stage 1 is the same as in T1: $E^{T2} = E^{T1}$. If $\lambda_1 > \frac{1}{2}$, then (x,t) is an SPE if and only if x is a NE of the following game of identical-interest:

P1 \ P2	X		Y	
X	$2b+c_1+c_2$	$2b+c_1+c_2$	$b+c_2$	$b+c_2$
Y	$b+c_1$	$b+c_1$	$2a$	$2a$

and the transfer function t in SPE satisfies

$$t_1(x) - t_2(x) = \frac{g_1(x) - g_2(x)}{2} \text{ for every } x.$$

Hence,

$$E^{T2} = \begin{cases} \{(X,X),(Y,Y)\} & \text{if } 2a \ge b+c_1, \\ \{(X,X)\} & \text{if } 2a < b+c_1. \end{cases}$$

A.4 Instructions

For those who will participate in the experiment

The experiment we are about to conduct will be used as academic material to analyze your choice behavior. You will be asked to make choices according to the following rules. Through those choices, you will earn points. All the points you earn will be converted to cash at a rate of 1.3 yen per point, which will be paid to you at the end of the experiment. All instructions for making the choices will be shown on the computer screen in front of you. The experiment is divided into four parts.

Part 1

Explanation of the experiment

Once the experiment begins, the computer will first randomly choose half of the participants as selectors. In Part 1, only these selectors will make choices. There is nothing for those who were not chosen as a selector to do in Part 1 of the experiment; if you were not chosen, please wait until Part 2 begins.

Those who were chosen as selectors will be asked to make six rounds of selections, as explained below. First, a selector will be paired with an individual who was not chosen as a selector. A selector will not be paired with another selector. Likewise, individuals who were not chosen as selectors will not be paired together. The individual you are going to be paired with will be randomly determined by the computer for each round of selection.

The selector will choose one of four options, A, B, C, and D. Based on the choice made by the selector, points are determined for you (the selector) and your counterpart (an individual who was not chosen as a selector).

Options	Your Points	Points for Your Counterpart
A	Numerical value *a*	Numerical value *b*
B	Numerical value *c*	Numerical value *d*
C	Numerical value *e*	Numerical value *f*
D	Numerical value *g*	Numerical value *h*

The table above shows the four options and their corresponding points for you and your counterpart. During the experiment, *a* through *h* will have specific numerical values and will change in each round.

For example, when, as the selector, you choose B, you will the receive points shown for the numerical value *c*, while your counterpart (an individual who was not chosen as a selector) will receive the points shown for numerical value *d*.

Formulae to Derive Points

Points for you (the selector) and your counterpart (an individual who was not chosen as a selector) are calculated according to the following formulae.

Your points $= 100 \times M \times N + 60 \times (1 - M) + s \times (1 - N)$

Points for your counterpart $= 100 \times M \times N + 60 \times (1 - N) + t \times (1 - M)$

M and N will each have a value of either 0 or 1; their combinations correspond to the options as follows:

Option A in the table above: M = 0 and N = 0

Option B in the table above: M = 0 and N = 1

Option C in the table above: M = 1 and N = 0

Option D in the table above: M = 1 and N = 1

The values of *s* and *t* are different and change in each round.

Numerical Value Example 1

Let us say the value of *s* is 30 and the value of *t* is 100. In this case, points will be given as follows:

When you choose Option A (M = 0, N = 0),

Your points = 100 × 0 × 0 + 60 × (1 - 0) + 30 × (1 - 0) = 90

Points for your counterpart = 100 × 0 × 0 + 60 × (1 - 0) + 100 × (1 - 0) = 160.

When you choose Option B (M = 0, N = 1),

Your points = 100 × 0 × 1 + 60 × (1 - 0) + 30 × (1 - 1) = 60

Points for your counterpart = 100 × 0 × 1 + 60 × (1 - 1) + 100 × (1 - 0) = 100.

When you choose Option C (M = 1, N = 0),

Your points = 100 × 1 × 0 + 60 × (1 - 1) + 30 × (1 - 0) = 30

Points for your counterpart = 100 × 1 × 0 + 60 × (1 - 0) + 100 × (1 - 1) = 60.

When you choose Option D (M = 1, N = 1),

Your points = 100 × 1 × 1 + 60 × (1 - 1) + 30 × (1 - 1) = 100

Points for your counterpart = 100 × 1 × 1 + 60 × (1 - 1) + 100 × (1 - 1) = 100.

In other words, the point table in this example will be as follows.

Options	Your Points	Points for Your Counterpart

A	90	160
B	60	100
C	30	60
D	100	100

Numerical Value Example 2

Let's say the value of *s* is 140 and the value of *t* is 55. In this case, the points will be as follows:

When you choose Option A (M = 0, N = 0),

Your points = 100 × 0 × 0 + 60 × (1 - 0) + 140 × (1 - 0) = 200

Points for your counterpart = 100 × 0 × 0 + 60 × (1 - 0) + 55 × (1 - 0) = 115.

When you choose Option B (M = 0, N = 1),

Your points = 100 × 0 × 1 + 60 × (1 - 0) + 140 × (1 - 1) = 60

Points for your counterpart = 100 × 0 × 1 + 60 × (1 - 1) + 55 × (1 - 0) = 55.

When you choose Option C (M = 1, N = 0),

Your points = 100 × 1 × 0 + 60 × (1 - 1) + 140 × (1 - 0) = 140

Points for your counterpart = 100 × 1 × 0 + 60 × (1 - 0) + 55 × (1 - 1) = 60.

When you choose Option D (M = 1, N = 1),

Your points = 100 × 1 × 1 + 60 × (1 - 1) + 140 × (1 - 1) = 100

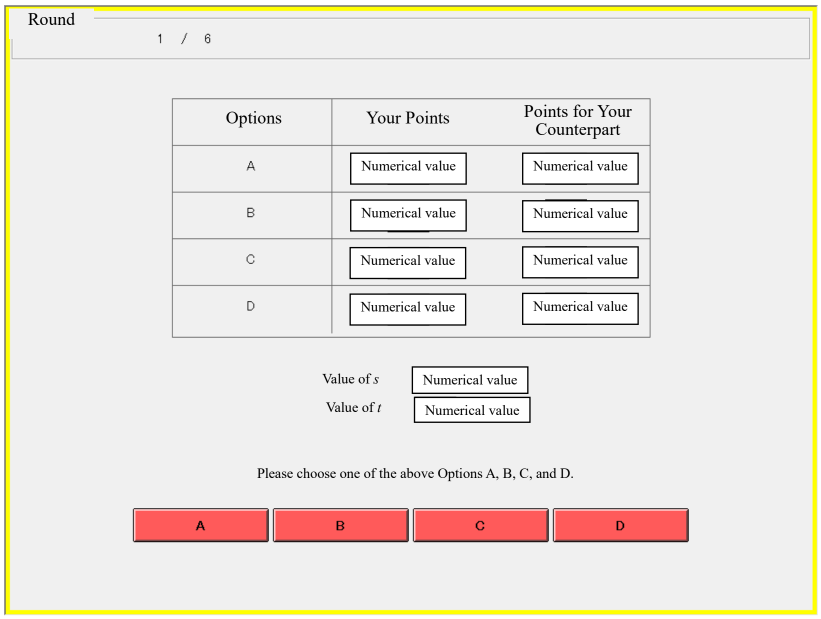
Points for your counterpart = 100 × 1 × 1 + 60 × (1 - 1) + 55 × (1 - 1) = 100.

The point table in this example will be as follows.

Options	Your Points	Points for Your Counterpart
A	200	115
B	60	55
C	140	60
D	100	100

Explanation of the Computer Screens and Tasks to Be Performed During the Experiment

Screen for the selector



The above screen will be displayed for those who are chosen as selectors. The selection round is shown at the top of the screen. In the above example, the round is the first of six.

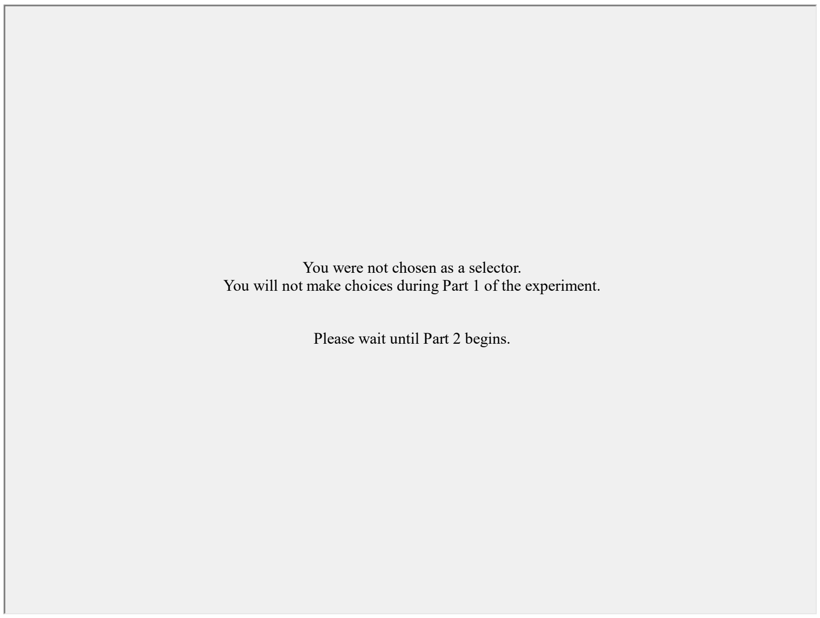
Four options and their corresponding points for you and your counterpart are shown in the center of the screen. These points will change each round. Below that, you will find the values of s and t that are used in the formulae to derive the points. Below that are the selection buttons.

First, please choose one option from A, B, C, and D and write it under "Your Choice" on the record sheet. Please also write why you made that choice in the "Reasons for Your Choice" section.

When you've finished writing, please click on the button for the option you wrote in "Your Choice."

Once all selectors click on an option button, the next round will begin. This completes one round of selection. This is repeated six times.

Screen for those who were not chosen as a selector



The above screen will be displayed for those who were not chosen as selectors. Please wait until Part 2 begins, as you will not make choices in Part 1 of the experiment.

You will now have three minutes to review the details of the experiment. If you have a question, please raise your hand quietly and the experimenter will answer you one-to-one. Please note that you are not allowed to communicate with the other participants.

Part 2

Explanation of the Experiment Details

Everyone will be asked to make choices starting in Part 2. You will make six rounds of selections as explained below. First, you will be paired with another participant. The individual you are paired with will be randomly chosen by the computer for each round of selection.

You will choose one of two options, X and Y. The points for you and your counterpart will be determined on the basis of the choices you and your counterpart make.

		Choice of Your Counterpart: X	Choice of Your Counterpart: Y
Your Choice	X	Numerical value *a*, Numerical value *b*	Numerical value *c*, Numerical value *d*
	Y	Numerical value *e*, Numerical value *b*	Numerical value *g*, Numerical value *h*

The X and Y on the left side represent your choice and the X and Y on the top represent the choice of your counterpart. The two numerical values in each cell represent the points for you and your counterpart corresponding to each combination of choices. The value on the left in each cell is your points and the value on the right is the points for your counterpart. During the experiment, numerical points *a* through *h* will have specific values and will change in each round.

For example, when your choice is X and the choice of your counterpart is Y, you will receive the points shown by numerical value *c* and your counterpart will receive the points shown by numerical value *d*.

Formulae to Derive Points

Your choices are 0 (X) and 1 (Y). Points for you and your counterpart are calculated according to the following formulae.

Your points $=$ 100 × (your choice) × (choice of your counterpart)
$+$ 60 × (1 $-$ your choice) $+$ s × (1 $-$ choice of your counterpart)

Points for your counterpart $=$ 100 × (your choice) × (choice of your counterpart)
$+$ 60 × (1 $-$ choice of your counterpart) $+$ t × (1 $-$ your choice)

The values of s and t are different and change in each round.

Numerical Value Example 1

Let us say the value of s is 30 and the value of t is 100. In this case, the points will be as follows:

When you choose 0 (X) and your counterpart chooses 0 (X),

Your points $= 100 \times 0 \times 0 + 60 \times (1 - 0) + 30 \times (1 - 0) = 90$

Points for your counterpart$= 100 \times 0 \times 0 + 60 \times (1 - 0) + 100 \times (1 - 0) = 160$

When you choose 0 (X) and your counterpart chooses 1 (Y),

Your points $= 100 \times 0 \times 1 + 60 \times (1 - 0) + 30 \times (1 - 1) = 60$

Points for your counterpart $= 100 \times 0 \times 1 + 60 \times (1 - 1) + 100 \times (1 - 0) = 100$

When you choose 1 (Y) and your counterpart chooses 0 (X),

Your points $= 100 \times 1 \times 0 + 60 \times (1 - 1) + 30 \times (1 - 0) = 30$

Points for your counterpart $= 100 \times 1 \times 0 + 60 \times (1 - 0) + 100 \times (1 - 1) = 60$

When you choose 1 (Y) and your counterpart chooses 1 (Y),

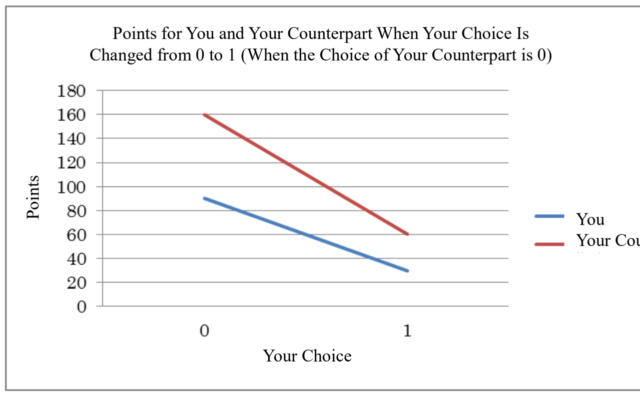
Your points $= 100 \times 1 \times 1 + 60 \times (1 - 1) + 30 \times (1 - 1) = 100$

Points for your counterpart $= 100 \times 1 \times 1 + 60 \times (1 - 1) + 100 \times (1 - 1) = 100$

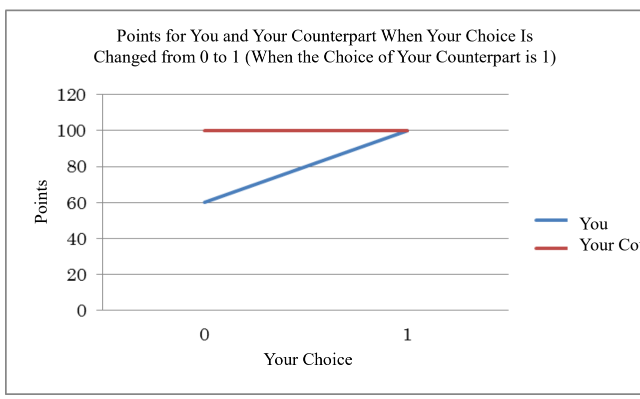
The point table in this example will be as follows.

		Choice of Your Counterpart	
		X	Y
Your Choice	X	90, 160	60, 100
	Y	30, 60	100, 100

When the choice of your counterpart is 0 (X) and your choice is changed from 0 to 1, the points for you and your counterpart can be graphed as follows.



When the choice of your counterpart is 1 (Y) and your choice is changed from 0 (X) to 1 (Y), the points for you and your counterpart can be graphed as follows.



Numerical Value Example 2

Let us say the value of s is 140 and the value of t is 55. In this case, the points will be as follows:

When you choose 0 (X) and your counterpart chooses 0 (X),

Your points $= 100 \times 0 \times 0 + 60 \times (1 - 0) + 140 \times (1 - 0) = 200$

Points for your counterpart $= 100 \times 0 \times 0 + 60 \times (1 - 0) + 55 \times (1 - 0) = 115$

When you choose 0 (X) and your counterpart chooses 1 (Y),

Your points $= 100 \times 0 \times 1 + 60 \times (1 - 0) + 140 \times (1 - 1) = 60$

Points for your counterpart $= 100 \times 0 \times 1 + 60 \times (1 - 1) + 55 \times (1 - 0) = 55$

When you choose 1 (Y) and your counterpart chooses 0 (X),

Your points $= 100 \times 1 \times 0 + 60 \times (1 - 1) + 140 \times (1 - 0) = 140$

Points for your counterpart $= 100 \times 1 \times 0 + 60 \times (1 - 0) + 55 \times (1 - 1) = 60$

When you choose 1 (Y) and your counterpart chooses 1 (Y),

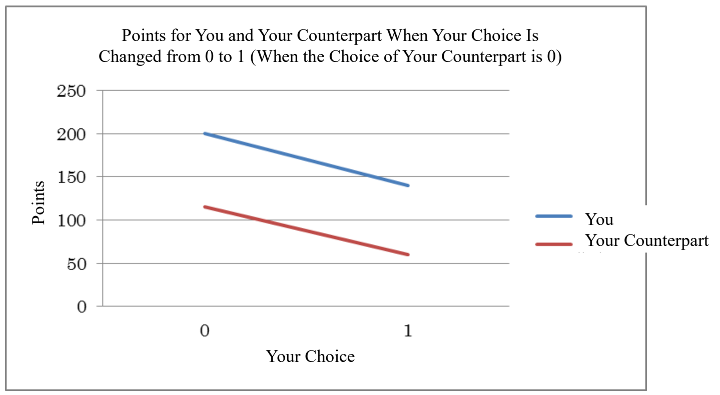
Your points $= 100 \times 1 \times 1 + 60 \times (1 - 1) + 140 \times (1 - 1) = 100$

Points for your counterpart $= 100 \times 1 \times 1 + 60 \times (1 - 1) + 55 \times (1 - 1) = 100$

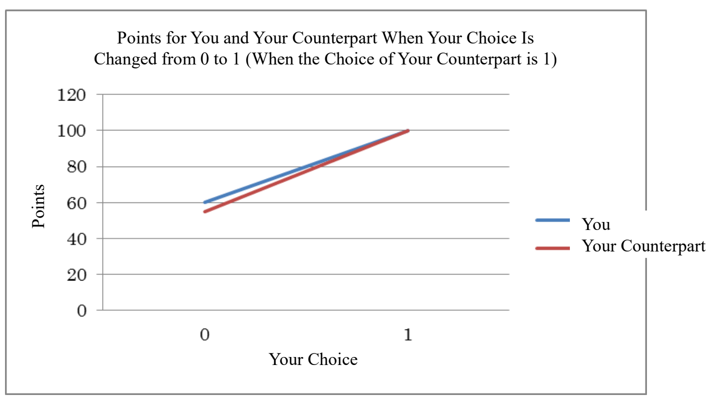
In other words, the point table in this example will be as follows.

		Choice of Your Counterpart	
		X	Y
Your Choice	X	200, 115	60, 55
	Y	140, 60	100, 100

When the choice of your counterpart is 0 (X) and your choice is changed from 0 to 1, the points for you and your counterpart can be graphed as follows.

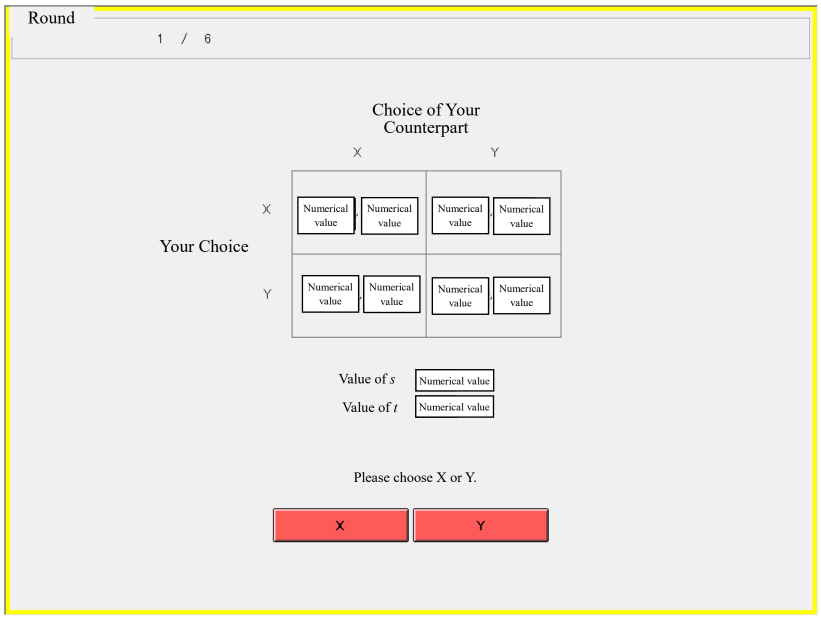


When the choice of your counterpart is 1 (Y) and your choice is changed from 0 (X) to 1 (Y), the points for you and your counterpart can be graphed as follows.



Explanation of the Computer Screen and the Tasks to Be Performed during the Experiment

Selection screen



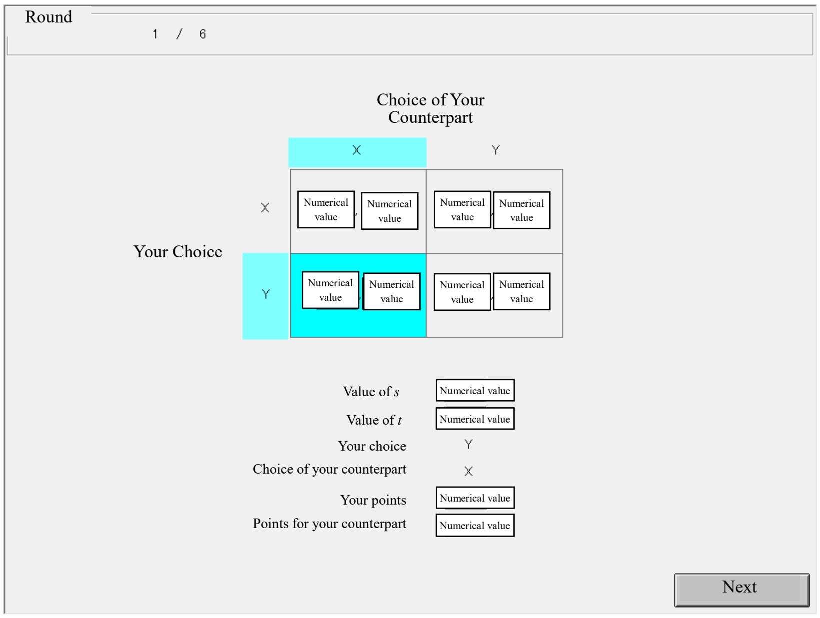
The selection screen will display the above. The point table is displayed in the center of the screen. Below it, you will find the values of *s* and *t* that are used in the formulae to derive the points. Below that are the selection buttons.

First, please choose X or Y and write it under "Your Choice" on the record sheet. Please also write why you made that choice in the "Reasons for Your Choice" section.

When you finish writing, please click on the button for the option you wrote in "Your Choice."

Once everyone clicks a selection button, the next round will begin.

Results screen



The results screen will display the above. The values of *s* and *t* that are used in the formulae to derive the points, your choice, the choice of your counterpart, your points, and the points for your counterpart are displayed below the center of the screen.

First, please write the choice of your counterpart displayed on the screen under “Choice of Your Counterpart” on the record sheet. Please also write your points displayed on the screen in the “Your Points” section.

When you finish writing, please click on the Next button at the lower right of the screen.

Once all subjects click the Next button, the next round will begin. This will complete the first round of selection. This process is repeated six times.

You will now have three minutes to review the details of the experiment. If you have a question, please raise your hand quietly and the experimenter will answer you one-to-one. Please note that you are not allowed to communicate with the other participants.

Part 3

Explanation of the Experiment Details

You will make six rounds of selections as explained below. First, you will be paired with another participant. The individual you are paired with will be randomly chosen by the computer for each round of selection. This part is composed of two stages.

Stage 1

You will choose one of two options, X and Y. The points for you and your counterpart will be determined on the basis of the choices you and your counterpart make.

		Choice of Your Counterpart: X	Choice of Your Counterpart: Y
Your Choice	X	Numerical value *a*, Numerical value *b*	Numerical value *c*, Numerical value *d*
	Y	Numerical value *e*, Numerical value *f*	Numerical value *g*, Numerical value *h*

This point table is read the same way as the previous one. Points *a* through *h* will each have a specific numerical value during the experiment and will change in each round. The numerical values in the point table are derived by the method explained in Part 2.

Stage 2

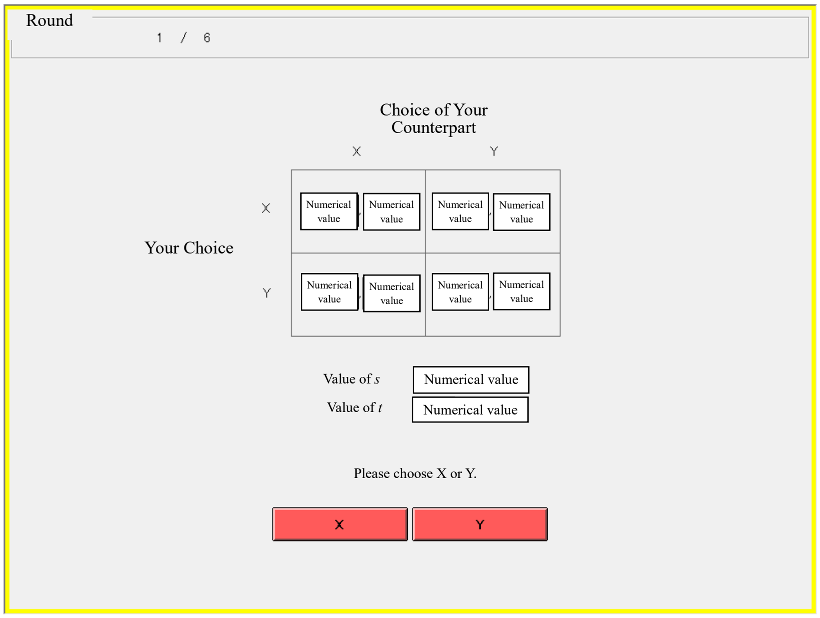
After verifying each other's choice and points in Stage 1, you will have an opportunity to give some of your points to your counterpart.

For example, let us say that your choice was X and the choice of your counterpart was Y in Stage 1 and that you and your counterpart earned *c* points and *d* points, respectively. Furthermore, let us say that you and your counterpart decided to give 10 points and 15 points to each other, respectively, in Stage 2. In this case, your final points will be $c - 10 + 15$, while the points for your counterpart will be $d - 15 + 10$.

Note: please choose an integer of 0 or larger that is smaller than the number of points you earned in Stage 1 to give to your counterpart.

Explanation of the Computer Screen and the Tasks to Be Performed during the Experiment

Selection screen in Stage 1



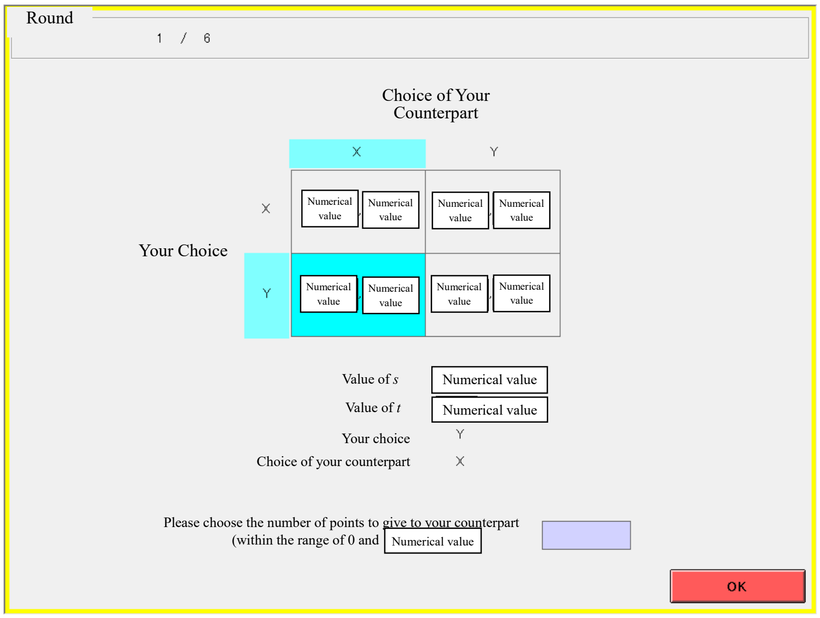
The selection screen in Stage 1 will display the above. The point table is displayed in the center of the screen. Below it, you will find the values of *s* and *t* that are used in the formulae to derive the points. Below that are the selection buttons.

First, please choose X or Y and write it under "Your Choice in Stage 1" on the record sheet. Please also write why you made that choice in the "Reasons for Your Choice in Stage 1" section.

When you finish writing, please click on the button for the option you wrote in "Your Choice in Stage 1."

Once all subjects click on a selection button, you will proceed to the selection screen in Stage 2.

Selection screen in Stage 2



The selection screen in Stage 2 will display the above. The values of s and t that are used in the formulae to derive the points as well as your choice and the choice of your counterpart in Stage 1 are displayed in the center of the screen. Below that is the selection box.

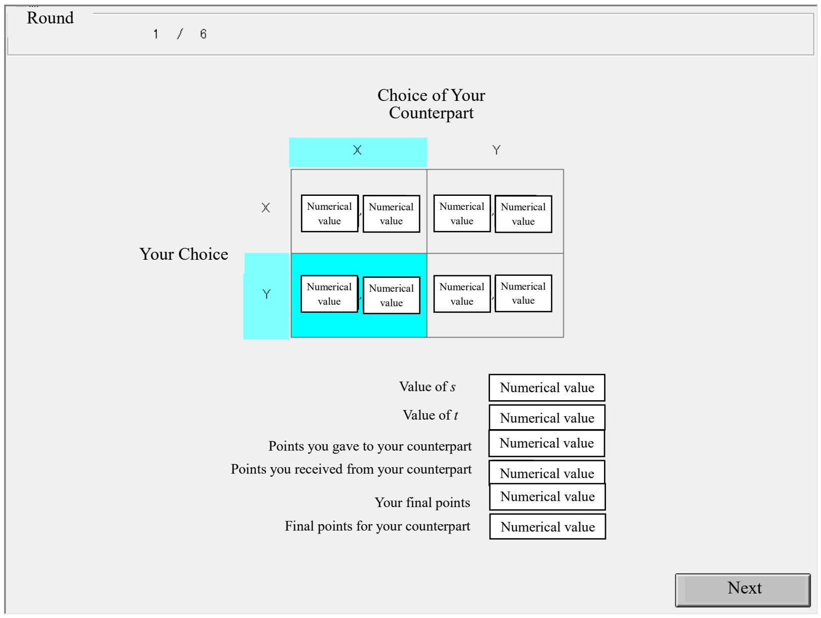
First, please write the choice of your counterpart displayed on the screen under “Choice of Your Counterpart in Stage 1” on the record sheet.

Next, please write the number of points you are giving to your counterpart under “Your Choice in Stage 2” on the record sheet. For the number of points to give, please choose an integer that is 0 or larger that does not exceed the number of points you earned in Stage 1. Please also write why you made that choice in the “Reasons for Your Choice in Stage 2” section.

When you finish writing, please enter the numerical value you wrote in “Your Choice in Stage 2” and click the OK button on the lower right.

Once all subjects click on the OK button, you will proceed to the results screen.

Results screen



The results screen will display the above. The values of *s* and *t* that are used in the formulae to derive the points as well as points you gave to your counterpart, points you received from your counterpart, your final points, and the final points for your counterpart are displayed in the center of the screen.

First, please write the points you received from your counterpart as displayed on the screen under "Choice of Your Counterpart in Stage 2" on the record sheet. Please also write your final points displayed on the screen in the "Your Point" section in the record sheet.

When you finish writing, please click the Next button on the lower right of the screen.
Once all subjects click on the Next button, the next round will begin. This will complete the first round of selection. This process is repeated six times.

You will now have three minutes to review the details of the experiment. If you have a question, please raise your hand quietly and the experimenter will answer you one-to-one. Please note that you are not allowed to communicate with the other participants.