CARF Working Paper

CARF-F-433

Accuracy and Retaliation in Repeated Games with Imperfect Private Monitoring: Experiments

Yutaka Kayaba
University of Tokyo

Hitoshi Matsushima
University of Tokyo

Tomohisa Toyama
International Christian University

March 18, 2016
Current Version: February 1, 2019

CARF is presently supported by The Dai-ichi Life Insurance Company, Limited, Nomura Holdings, Inc., Sumitomo Mitsui Banking Corporation, MUFG Bank, The Norinchukin Bank, The University of Tokyo Edge Capital Co., Ltd., Finatext Ltd., Sompo Holdings, Inc., Tokio Marine & Nichido Fire Insurance Co., Ltd. and Nippon Life Insurance Company. This financial support enables us to issue CARF Working Papers.

Accuracy and Retaliation in Repeated Games with Imperfect Private Monitoring: Experiments[1]

Yutaka Kayaba[2]
University of Tokyo

Hitoshi Matsushima[3]
University of Tokyo

Tomohisa Toyama[4]
International Christian University

March 18, 2016
Current Version: February 1, 2019

Abstract

We experimentally examine repeated prisoner's dilemma with random termination, in which monitoring is imperfect and private. Our estimation indicates that a significant proportion of subjects follow generous tit-for-tat strategies, which are stochastic extensions of tit-for-tat. However, the observed retaliating policies are inconsistent with the generous tit-for-tat equilibrium behavior. Contrary to the prediction of the equilibrium theory, subjects tend to retaliate more with high accuracy than with low accuracy. They tend to retaliate more than the equilibrium theory predicts with high accuracy, while they tend to retaliate less with low accuracy.
JEL Classification Numbers: C70, C71, C72, C73, D03.
Keywords: Repeated Prisoner's Dilemma, Imperfect Private Monitoring, Experiments, Generous Tit-for-Tat, Retaliation Intensity.

---

[1] Matsushima, Toyama, and Yagi (2007); Matsushima and Toyama (2011); and Matsushima, Tanaka, and Toyama (2013) are earlier versions of this paper. Research for this study was financially supported by a grant-in-aid for scientific research (KAKENHI 25285059) from the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of the Japanese government, as well as by the Center of Advanced Research in Finance at the University of Tokyo. We are grateful to the editor and the anonymous referees for their helpful comments and suggestions. We are grateful to Hide Ichimura and the participants of seminars at the SAET conference, Hitotsubashi University, and the University of Tokyo for their comments. We appreciate the helpful support of Tomomi Tanaka and Nobuyuki Yagi. All errors are our own.
[2] Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: ykayaba@e.u-tokyo.ac.jp
[3] Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: hitoshi@e.u-tokyo.ac.jp
[4] Division of Arts and Sciences, College of Liberal Arts, International Christian University, 3-10-2 Osawa, Mitaka-shi, Tokyo 181-8585, Japan. E-mail: tomohisa.toyama@gmail.com

1. Introduction

It is widely accepted that long-run strategic interaction facilitates collusion among players whose interest conflict with each other. The premise is that each player observes information about the actions that opponents have selected previously. However, even if the monitoring of opponents' actions is *imperfect* (i.e., each player cannot directly observe the opponents' action choices but can observe informative signals), theoretical studies have shown that sufficiently patient players can still employ, to a greater or lesser degree, cooperative strategies as an equilibrium. To be more precise, the folk theorem generally indicates that if the discount factor is sufficient (i.e., close to unity) and each player can indirectly—but not directly—observe opponents' action choices through noisy signals, a wide variety of allocations can be attained by subgame perfect equilibria in the infinitely repeated game (e.g., Fudenberg, Levine, and Maskin, 1994; Sugaya, 2012). Indeed, the folk theorem is applicable to a very wide range of strategic conflicts.

However, the folk theorem does not specify what kind of equilibria emerge empirically or the strategies people follow that are actually associated with the equilibria. Given the lack of consensus on the strategies people empirically follow, this study experimentally analyses our subjects' behavior in a repeated prisoner's dilemma.

Our experimental setup is imperfect monitoring. Each player cannot directly observe her opponent's action choice; instead, she observes a signal, which is either *good* or *bad*. The good (bad) signal is more likely to occur when the opponent selects the cooperative (defective) action rather than the defective (cooperative) action. The probability that a player observes the good (bad) signal when the opponent selects the cooperative (defective) action is referred to as monitoring accuracy; it is denoted by $p \in (\frac{1}{2}, 1)$. The study experimentally controls the levels of monitoring accuracy as treatments (high accuracy $p = 0.9$ and low accuracy $p = 0.6$). Specifically, the monitoring technology is *private* in that each player cannot receive any information about what the opponent observes about the player's choices (i.e., the signals are observable only by the receivers).

To examine the prevalence of the strategies our subjects employ, we employ the strategy frequency estimation method (SFEM) developed by Dal Bó and Fréchette (2011). SFEM lists various potential strategies, such as tit-for-tat (TFT), grim-trigger, lenience, and long-term punishment strategies; all these comprise a significant proportion of strategies found in existing studies of experimental repeated games. SFEM then estimates the frequencies of each strategy, where the heterogeneity of strategies people follow is treated explicitly. There are existing experimental studies using SFEM to examine the prevalence of strategies (e.g., Fudenberg, Rand, and Dreber, 2012; Aoyagi, Bhaskar, and Fréchette, 2019).

Unlike these experimental studies, we rigorously include stochastic strategies in our SFEM list. Importantly, we include straightforward stochastic extensions of TFT, that is, *generous tit-for-tat* (g-TFT). According to TFT, a player mimics her opponent’s action by making the cooperative (defective) action choice whenever she observes the good (bad) signal. G-TFT is defined as its stochastic extension, according to which, a player makes the cooperative (defective) action choice with a higher probability when she observes the good signal than when she observes the bad signal.

TFT is regarded as a reciprocal behavioral mode that describes cooperation, retaliation, and forgiveness in a simple and tractable manner. [5] However, TFT has several drawbacks. For instance, TFT generally fails to be a subgame perfect equilibrium. Further, TFT cannot escape the death spiral, where players endlessly repeat the alternating play of cooperation and defection, once it falls into it.

In contrast, g-TFT overcomes these drawbacks; G-TFT equilibria always exist, irrespective of the level of monitoring accuracy, provided the discount factor is sufficient. G-TFT can avoid the death spiral of endless retaliations among players. Hence, it would be reasonable to expect that actual human beings and animals conduct such random experimentations as g-TFT implies. Indeed, evolutionary biology finds that animals maintain peaceful coexistence, instead of weak cost, by adopting g-TFT (e.g., Molander, 1985; Nowak and Sigmund, 1992). In human societies, g-TFT is expected to provide an opportunity to avoid the crisis of nuclear war and build a peaceful relationship. G-TFT strategies offer the briefest way to explain the principle of cooperation.

In this study, using the SFEM framework, we estimate the fraction of our subjects

---

[5] Axelrod (1984) showed that TFT is one of the most successful strategies in round-robin tournament experiments and computer simulations. In response, the public praised TFT as the basic principle of implicit collusion. On the other hand, some game theorists criticize the insufficiency of the Axelrod's evolutionary simulations (e.g., Binmore, 1994, Chapter 3).

playing g-TFT strategies, as well as those playing other strategies that frequently appear in existing studies of experimental repeated games. We also include various long-memory strategies and their stochastic versions in the list.

Our estimates indicate that a significant proportion (about 70%) of our subjects follow g-TFT strategies, albeit heterogeneous ones. Although existing empirical or theoretical studies stress grim, lenient, and long-term punishing strategies, our experimental results demonstrate that the empirical importance of g-TFT is supported experimentally as well as the theoretical importance mentioned earlier.

Moreover, observing that many of our subjects follow g-TFT strategies, we empirically examine their retaliation policies. We focus on the contrast in the probabilities of cooperative action choices contingent on good and bad signals, that is referred as the *retaliation intensity*. [6]

Fixing a sufficient discount factor, the retaliation intensities are common across all g-TFT equilibria, depending on the level of monitoring accuracy. This common retaliation intensity is decreasing at the level of monitoring accuracy. This decreasing property plays the central role in making use of the improvement of monitoring technology and effectively saving the welfare loss caused by the monitoring imperfection.

However, the retaliation intensities observed in our experimental data contradict the predictions of the above-mentioned equilibrium theory; our subjects tend to retaliate more in the high accuracy treatment than in the low accuracy treatment. They tend to retaliate with an intensity that is greater than the level implied by the g-TFT equilibria in the high accuracy treatment, while they tend to retaliate less than expected in the low accuracy treatment. Hence, when the monitoring is accurate, the expected payoff to an individual subject from cooperation tends to be much greater than that from defection; when the monitoring is inaccurate, the expected payoff from cooperation tends to be much less than that from defection.

The rest of this paper is organized as follows. Section 2 reviews the literature. Section 3 presents the basic model. Section 4 introduces the g-TFT strategy. Section 5 explains the experimental design. Section 6 presents the experimental results for aggregate behavior. Section 7 explains the SFEM. Section 8 presents the experimental

[6] Note that TFT corresponds to the g-TFT whose retaliation intensity equals unity.

results for individual strategies.

Section 9 discusses an alternative experimental design. The strategies that subjects actually follow could depend on various environmental factors, such as the kind of interpretation about infinite interactions that is adopted, how it is described in the experimental instruction, how many times a subject engages in tasks with a single experimental treatment, and so on. We conduct additional experiments to explore how the other experimental design could alter the choice of strategies, and report the results. Section 10 concludes.

2. Literature Review

This study contributes to the long history of research in the repeated game literature. Equilibrium theory demonstrates folk theorems in various environments, which commonly show that a wide variety of outcomes is sustained by perfect equilibria, provided the discount factor is sufficient. Fudenberg and Maskin (1986) and Fudenberg, Levine, and Maskin (1994) proved folk theorems for perfect monitoring and imperfect public monitoring, respectively. These studies utilized the self-generation nature of perfect equilibria explored by Abreu (1988) and Abreu, Pearce, and Stacchetti (1990), which, however, crucially relied on the publicity of signal observations.

For the study of imperfect private monitoring, Ely and Välimäki (2002) and Piccione (2002) explored the belief-free nature as an alternative to self-generation, which motivates a player to select both cooperative action and defective action at all times. These studies presented the folk theorem for prisoner's dilemma, wherein monitoring is private and almost perfect. [7] Based on this belief-free nature, Molander (1985), Nowak and Sigmund (1992), and Takahashi (2010) studied g-TFT strategies in various situations, such as biological populations and large communities with random matching. Matsushima (2013) studied g-TFT equilibria in a class of prisoner's dilemma games in

[7] For a survey of almost perfect private monitoring, see Mailath and Samuelson (2006). Matsushima (2004) proved the folk theorem in the prisoner's dilemma game with imperfect private monitoring by constructing review strategy equilibria as lenient behavior with long-term punishments, in which we permit the monitoring technology to be arbitrarily inaccurate. Sugaya (2012) proved the folk theorem with imperfect private monitoring for a very general class of infinitely repeated games by extending self-generation to imperfect private monitoring and then combining it with the belief-free nature.

which monitoring is private and far from perfect.

The literature of experimental studies on repeated games has examined the determinants of cooperation and tested various theoretical predictions to find clues to resolving the multiplicity problem (for a review, see Dal Bó and Fréchette, 2016). The SFEM, which is employed in this study, is frequently used in the literature on experimental repeated games (e.g., Dal Bó and Fréchette, 2011; Fudenberg, Rand, and Dreber, 2012; Aoyagi, Bhaskar, and Fréchette, 2019; Breitmoser, 2015). This study includes various stochastic strategies in our SFEM list. The inclusion of such stochastic action choices is scant in the literature of experimental repeated games. Fudenberg, Rand, and Dreber (2012) include only a few g-TFT strategies, aiming only to perform robustness checks for their claim that their experimental subjects tend to be lenient. [8]

By contrast, we rigorously include many variants of g-TFT in our SFEM list. We also include various long-memory strategies and their stochastic variants, which are more complicated than TFT and g-TFT. Our experimental results support the idea that players do not retaliate every time they observe a single occurrence of a bad signal, not because they delay punishment until additional occurrence of bad signals, but because they employ (non-trivial) stochastic strategies. This finding contrasts with that of Fudenberg, Rand, and Dreber (2012), and Aoyagi, Bhaskar, and Fréchette (2019). Both studies stress long-memory strategies rather than memory-one strategies within the scope of deterministic strategies.

3. The Model

This study investigates a repeated game played by two players (players 1 and 2), using a discrete time horizon. This game has a finite round-length, but the terminating round is randomly determined, and, therefore, unknown to players. The component game of this repeated game is denoted by $(S_i, u_i)_{i\in\{1,2\}}$, where S_i denotes the set of all actions for player $i \in \{1,2\}$, $s_i \in S_i$, $S \equiv S_1 \times S_2$, $s \equiv (s_1, s_2) \in S$, $u_i : S \to R$, and $u_i(s)$

[8] The experimental setup in Fudenberg, Rand, and Dreber (2012) is imperfect public monitoring in which the stochastic strategies have relatively less importance than the case of imperfect private monitoring.

denotes the payoff for player i induced by action profile $s \in S$.

In each round, two noisy signals $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$ occur after the action choices are made, where Ω_i denotes the set of possible ω_i, $\omega \equiv (\omega_1, \omega_2)$, and $\Omega \equiv \Omega_1 \times \Omega_2$. A signal profile $\omega \in \Omega$ is randomly determined according to a conditional probability function $f(\cdot \mid s): \Omega \to R_+$, where $\sum_{\omega \in \Omega} f(\omega \mid s) = 1$ for all $s \in S$. We assume *full support* in that $f(\omega \mid s) > 0$ for all $\omega \in \Omega$ and $s \in S$.

Let $f_i(\omega_i \mid s) \equiv \sum_{\omega_j \in \Omega_j} f(\omega \mid s)$, and we assume that $f_i(\omega_i \mid s)$ is independent of s_j. Hence, we denote $f_i(\omega_i \mid s_i)$ instead of $f_i(\omega_i \mid s)$. We utilize $\omega_i \in \Omega_i$ to denote *the signal for player i's action choice*. Player i's action choice s_i influences the occurrence of the signal for her action choice ω_i, but does not influence the occurrence of the signal for the opponent's action choice ω_j, where $j \neq i$.

We assume that monitoring is *imperfect*. In every round $t \in \{1, 2, ...\}$, each player i cannot directly observe either the action $s_j(t) \in S_j$ that the opponent $j \neq i$ has selected, or the realized payoff profile $u(s(t)) = (u_1(s(t)), u_2(s(t))) \in R^2$, where we denote by $s(t) = (s_1(t), s_2(t)) \in S$ the action profile selected in round t. Instead, player i observes the signal for opponent j's action choice $\omega_j(t) \in \Omega_j$ through which player i monitors opponent j's action choice $s_j(t)$ indirectly and imperfectly. [9]

We further assume that monitoring is *private*. Each player cannot know what kind of signal her opponent receives about her own action choice. Hence, each player i knows $s_i(t)$ and $\omega_j(t)$ but does not know either $s_j(t)$ or $\omega_i(t)$.

This study specifies the component game as a *prisoner's dilemma with symmetry and additive separability* as follows: $S_1 = S_2 = \{C, D\}$; $u_1(C,C) = u_2(C,C) = 1$;

[9] Our specification of monitoring structure is in contrast to previous works, such as Green and Porter (1984) and Aoyagi and Fréchette (2009). These studies commonly assumed that the distribution of a noisy signal depends on all players' action choices, while we assume the abovementioned independence.

$u_1(D,D) = u_2(D,D) = 0$; $u_1(C,D) = u_2(D,C) = -g$; and $u_1(D,C) = u_2(C,D) = 1+g$, where we assume $g > 0$. Let us call C and D the *cooperative* action and *defective* action, respectively. Selecting C instead of D costs g but gives the opponent the benefit $1+g$. The payoff vector induced by the cooperative action profile (C,C) maximizes welfare $u_1(s)+u_2(s)$ with respect to $s \in S$. The defective action profile (D,D) is the dominant strategy profile and the unique Nash equilibrium.

We specify $\Omega_i = \{c,d\}$, $f_i(c\,|\,C) = f_i(d\,|\,D) = p$, and $\frac{1}{2} < p < 1$. Let us call c and d the *good signal* and *bad signal*, respectively. The probability index p implies the level of *monitoring accuracy*. The greater p is, the more accurately each player can monitor the opponent's action choice. Inequality $p > \frac{1}{2}$ implies that the probability of a good signal c occurring for a player is greater when this player selects C than when she selects D.

Let $h(t) = (s(\tau), \omega(\tau))_{\tau=1}^{t}$ denote the *history up to round* t. $H = \{h(t)\,|\,t = 0,1,...\}$ denotes the set of possible histories, where $h(0)$ denotes the null history. Player i's strategy in the repeated game is defined as $\sigma_i : H \to [0,1]$. According to σ_i, she selects cooperative action C with probability $\sigma_i(h(t-1))$ in each round t, provided history $h(t-1)$ up to round $t-1$ occurs. Let Σ_i denote the set of all strategies for player i. Let $\sigma \equiv (\sigma_1, \sigma_2)$ and $\Sigma \equiv \Sigma_1 \times \Sigma_2$.

We assume *constant random termination* in which $\delta \in (0,1)$ denotes the probability of the repeated game continuing after the end of each round t, provided this game continues up to round $t-1$. Hence, the repeated game is terminated at the end of each round $t \geq 1$ with probability $\delta^{t-1}(1-\delta)$. The expected payoff for player i induced by $\sigma \in \Sigma$, when the level of monitoring accuracy is given by $p \in (0,1)$ is defined as

$$U_i(\sigma; p) \equiv (1-\delta) E[\sum_{\tau=1}^{\infty} \delta^{\tau-1} u_i(s(\tau))\,|\,\sigma, p].$$

where $E[\cdot\,|\,\sigma, p]$ denotes the expectation operator conditional on (σ, p). A strategy profile $\sigma \in \Sigma$ is said to be *an equilibrium* in the repeated game with monitoring accuracy $p \in (0,1)$ if

$$U_i(\sigma; p) \geq U_i(\sigma_i', \sigma_j; p) \text{ for all } i \in \{1,2\} \text{ and all } \sigma_i' \in \Sigma_i .$$ [10]

For each history $h(t) \in H$ up to round t, we define the *frequency of cooperative action choice* C, or the *cooperation rate*, by

$$\rho(h(t)) \equiv \frac{\left|\{\tau \in \{1,...,t\} \mid S_1(\tau) = C\}\right| + \left|\{\tau \in \{1,...,t\} \mid S_2(\tau) = C\}\right|}{2t}.$$

The expected frequency of cooperative action choice C (i.e., the expected cooperation rate induced by $\sigma \in \Sigma$) is denoted by

$$\rho(\sigma; p) \equiv \frac{E[\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t\rho(h(t)) \mid \sigma, p]}{\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t}.$$

4. Generous Tit-For-Tat Strategy

A strategy $\sigma_i \in \Sigma_i$ is said to be *generous tit-for-tat (g-TFT)* if there exists $(q, r(c), r(d)) \in [0,1]^3$ such that $r(c) > 0$, $\sigma_i(h(0)) = q$, and for every $t \geq 2$ and $h(t-1) \in H$,

$$\sigma_i(h(t-1)) = r(c) \qquad \text{if } \omega_j(t-1) = c,$$

and

$$\sigma_i(h(t-1)) = r(d) \qquad \text{if } \omega_j(t-1) = d.$$

In round 1, player i makes the cooperative action choice C with probability q. In each round $t \geq 2$, player i makes the cooperative action choice C with probability $r(\omega_j)$ when she observes signal $\omega_j(t-1) = \omega_j$ for the opponent's action choice in round $t-1$. Hence, we simply write $(q, r(c), r(d))$ instead of σ_i for any g-TFT strategy. A g-TFT strategy $(q, r(c), r(d))$ is said to be an *equilibrium* in the repeated game with accuracy $p \in (0,1)$ if the corresponding symmetric g-TFT strategy profile is an

[10] The full-support assumption makes the distinction between Nash equilibrium and sequential equilibrium redundant.

equilibrium in the repeated game with accuracy p.

Define

$$w(p,g,\delta)=w(p)\equiv\frac{g}{\delta(2p-1)(1+g)}.$$

Note that $w(p)>0$, and

$$w(p)\leq 1\text{, if and only if } \delta\geq\frac{g}{(2p-1)(1+g)}.$$

According to the belief-free nature (e.g., Ely and Välimäki, 2002; Piccione, 2002; Bhaskar and Obara, 2002), we can prove that a g-TFT strategy $(q,r(c),r(d))$ is an equilibrium if and only if the difference in cooperation rate between the good and bad signals, that is, $r(c)-r(d)$, is equal to $w(p)$.

The Proposition: *A g-TFT strategy* $(q,r(c),r(d))$ *is an equilibrium in the repeated game with accuracy* p *if and only if* $0<w(p)\leq 1$*, and*

(1) $$r(c)-r(d)=w(p).$$

Proof: See Appendix A.

This study regards the observed difference in cooperation rate between the good and bad signals as the intensity with which subjects retaliate against their opponents. We call it *the retaliation intensity*. Note from this proposition that if subjects play a g-TFT equilibrium, then the resultant retaliation intensity should be (approximately) equal to $w(p)$.

Importantly, the retaliation intensity implied by the g-TFT equilibria $w(p)$ is *decreasing in* p; the less accurate the monitoring technology is, the more severely players retaliate against their opponents. This decreasing property is essential for understanding how players overcome the difficulty of achieving cooperation under imperfect private monitoring.

To incentivize a player to make the cooperative action choice, it is necessary that her opponent makes the defective action choice when observing the bad signal more often than when observing the good signal. In other words, *the retaliation intensity must be* *positive*.

When monitoring is inaccurate, it is difficult for the player's opponent to detect whether the player actually makes the cooperative action choice or the defective action choice. In this case, enhancement in retaliation intensity is necessary to incentivize the player. Hence, *the retaliation intensity must be decreasing at the level of monitoring accuracy*. This decreasing property plays a central role in improving welfare by utilizing noisy signals as much as possible. Since monitoring is imperfect, it is inevitable that the opponent observes the bad signal even if the player actually makes the cooperative action choice. This inevitably leads to welfare loss, because the opponent might retaliate against the player even if she has actually made the cooperative action choice. In this case, if the monitoring technology is more accurate, the opponent can incentivize the player well by being less sensitive to whether the observed signal is good or bad, thereby safely lowering the retaliation intensity. This serves to decrease the welfare loss caused by the monitoring imperfection. Hence, it is crucial from the viewpoint of welfare to examine whether the experimental results satisfy this decreasing property.

5. Experimental Design

We conducted four sessions of computer-based laboratory experiments[11] at the Center for Advanced Research for Finance, University of Tokyo, in October 2006. We recruited 108 subjects from a subject pool consisting of undergraduate and graduate students in various fields. Our subjects were given monetary incentives; the points earned in the experiments were converted into Japanese yen at a fixed rate (0.6 JPY per point). In addition, our subjects were each paid a fixed participation fee of 1,500 JPY.

To simplify the structure of the game, we adopt the prisoner's dilemma with symmetry and additive separability for our component game, where we assume $g = 2/9$. The payoff parameters have a structure in which the cost for cooperation, g, is small so that g-TFT equilibria exist even if the monitoring technology is poor. To make all payoffs greater than 0, we further make a positive linear transformation with variable coefficient

[11] The experiment was programmed and conducted with z-Tree software (Fischbacher, 2007).

45 and constant coefficient 15. The payoff matrix employed in the experiments is displayed in Table 1. The labels on the actions and signals are presented in neutral language (i.e., the actions are labeled "A" and "B" instead of "C [cooperation]" and "D [defection]," and the signals are labeled "a" and "b" instead of "c [good]" and "d [bad]").

[TABLE 1 HERE]

The experiments have two treatments that differ with respect to monitoring accuracy; in one treatment, monitoring accuracy is high. In this treatment, the signals the player observes and the action choices by the opponent coincide with a 90% chance ($p = 0.9$); the chance of mismatch is 10%. We refer to this treatment as the "high accuracy treatment." The other treatment is the case in which the monitoring technology is poorer. The chance of the signals observed by the player matching with the opponent's action choices is only 60% ($p = 0.6$). We refer to this treatment as the "low accuracy treatment."

All subjects participating in the four sessions received both treatments, but the treatment order was counterbalanced to minimize order effects. [12] The first two sessions (52 subjects) started with three repeated games of the low accuracy treatment and then proceeded to three repeated games of the high accuracy treatment. The second two sessions (56 subjects) were conducted in the reverse order, starting with the high accuracy treatment and ending with the low accuracy treatment. This within-subject design enables us to observe the subject-wise changes of retaliation intensities across the two treatments. Each treatment was preceded by a short repeated game, consisting of two rounds, to let the subjects understand the new treatment. Subjects were randomly paired at the start of each repeated game, and the pairs remained unchanged until the end of the repeated game. Table 2 displays the summary of the treatment order, number of subjects, and game lengths in repeated games that are determined by the continuation probability explained hereafter.

[TABLE 2 HERE]

[12] Nonetheless, order effects could occur in an individual level. We observed some order effects, which are documented in Online Appendix C.

We employed constant random termination. We let the continuation probability be $\delta = 0.967$ $(= 29/30)$; the very high continuation probability mimics the discount factor that is sufficiently large, thereby supporting the existence of equilibria in which players collude with each other. Hence, by assuming that the discount factor is close to unity and that the gain from deviation is small, we can make the incentive to cooperate compatible with the incentive to retaliate in terms of monetary interests, even if the monitoring technology is poor. In fact, there exists a cooperative g-TFT equilibrium even with low accuracy ($p = 0.6$).

The repeated game was terminated in the current round with probability $1-\delta\,(=1/30)$, and the subjects were re-matched with a new opponent and proceeded to play the next repeated game. Our subjects were not informed which round was the final round for each repeated game. We only announced a maximum length (98 rounds) in the experiments, which is sufficiently larger than the expected number of rounds computed from the terminating probability (30 rounds). To help our subjects understand that the termination rule is stochastic and the probability of termination in each round is 1/30, we presented 30 cells (numbered from 1 to 30) on the computer screen at the end of each round. One number was selected randomly at the end of each round, and the repeated game was terminated if the number 30 was selected by chance and the 30th cell in the computer screen turned green. Otherwise, all the cells numbered 1 to 29 turned green simultaneously and the repeated game continued. The screen is demonstrated in Online Appendix D.

Each subject was informed of the rules of the game, and the ways in which the game would proceed, with the aid of printed experimental instructions. The instructions were explained aloud by a recorded voice. Using the computer screen during the experiments, our subjects were able to review the structure of the game and the history up to the latest round; the history consisting of her own actions and the signals of the opponent's actions. See Online Appendix H for the experimental instructions and the images of the computer screen, which have been translated into English from the original Japanese. [13]

---

[13] In Section 9, we demonstrate an alternative experimental design, and present the difference of the results between the two experimental designs.

6. Aggregate Level Analysis of Experimental Results

6.1. Overall Cooperation Rates and Round 1 Cooperation Rates

[TABLE 3 HERE]

Table 3 displays the descriptive summary of the data. The entire group of 108 subjects made 8,864 decisions in the high accuracy treatment and 9,144 decisions in the low accuracy treatment. The overall frequency of cooperative choices (i.e., the cooperation rate) is 0.672 in the high accuracy treatment, and 0.355 in the low accuracy treatment. Statistically, the former frequency is significantly larger than the latter ($p < 0.001$, Wilcoxon matched-pair test for individual-level means). A first look at the cooperation rates suggests that our subjects cooperate more as the monitoring technology is improved.

[TABLE 4 HERE]

Table 4 presents the round 1 cooperation rates, that is, the frequency of cooperative action choices in round 1. The round 1 cooperation rate is 0.781 in the high accuracy treatment and 0.438 in the low accuracy treatment, with the latter frequency being significantly smaller ($p < 0.001$) than the former. Our subjects tend to start repeated games with cooperative action choices in the high accuracy treatment; however, their motivation for cooperation diminishes to such an extent with an increase in the noise in the signal that the chance that they start with cooperative actions becomes less than 50%.

Reacting to the change in the signal quality, they switch their strategies from cooperative ones in the high accuracy treatment to less cooperative ones in the low accuracy treatment.

6.2. Signal-Contingent Cooperation Rates

Table 4 also displays signal-contingent cooperation rates. The frequency of cooperative actions after observing a good signal is computed as the simple mean of all choices, and denoted by $r(c;p)$. The simple mean of all choices in the high accuracy treatment is 0.852. In addition, Table 4 reports an alternative value, which is the mean of individual-level means; it is concerned with the possibility that the behavior of the cooperative subjects might be over-represented in the simple mean of choices. [14] The mean of individual-level means in the high accuracy treatment is 0.788, which is 0.064 less than the simple mean of choices, implying that over-representation might exist. However, both measures are consistently high, reaching around 0.8, thereby indicating that our subjects are quite cooperative when they observe a good signal in the high accuracy treatment.

In the low accuracy treatment, the cooperation rate after observing a good signal is not as high as in the high accuracy treatment. The simple mean of the cooperative choices is 0.437, and the mean of individual-level means is 0.423. As in the case of the round 1 cooperation rate, our subjects are reluctant to cooperate even after observing good signals in the low accuracy treatment; this is not the case in the high accuracy treatment. The direct comparison of the cooperation rates between the two treatments indicates that the cooperation rate after observing a good signal in the high accuracy treatment is larger than that in the low accuracy treatment ($p < 0.001$ for both the simple mean and the mean of individual-level means).

As in the case of the cooperation rates after observing a bad signal, denoted by $r(d;p)$, the simple mean of cooperative choices in the high accuracy treatment is 0.344, and the mean of individual-level means is 0.443. Both measures are still below 50%, indicating that our subjects tend to defect rather than cooperate after observing a bad signal in the high accuracy treatment.

The tendency of our subjects to defect after observing a bad signal is more apparent in the low accuracy treatment. The simple mean of cooperative actions across all choices is 0.272 and the mean of individual-level means is 0.279, both of which are consistently

[14] Since subjects who are cooperative might observe good signals more often and adopt cooperative actions more often, their cooperative choices might be over-represented in the computation of cooperation rates that are contingent on a good signal.

smaller than the means in the high accuracy treatment ($p < 0.001$ for both means). Observing a bad signal, our subjects tend to defect more in the low accuracy treatment than in the high accuracy treatment.

The overall picture of the round 1 cooperation rates and the signal-contingent cooperation rates shown above robustly demonstrates that, irrespective of the type of signal observed, our subjects make more cooperative action choices when the signal quality is better.

RESULT 1-a: Our subjects tend to cooperate more in the high accuracy treatment than in the low accuracy treatment, adapting their strategies according to signal accuracy.

6.3. Retaliation Intensity

[TABLE 5 HERE]

We examine whether the observed retaliation intensity, defined as the difference in cooperation rate $r(c; p) - r(d; p)$, coincides with the theoretical value implied by the g-TFT equilibria $w(p)$. Table 5 presents the retaliation intensities at the aggregate level.

In the high accuracy treatment, the retaliation intensity $r(c;0.9) - r(d;0.9)$ is 0.508 in the simple mean of all choices and 0.352 in the mean of individual-level means. Both measures consistently differ from 0 in a statistically significant manner ($p < 0.001$ for both). These results indicate that our subjects use signal-contingent information in their action choices. Statistically, both measures are significantly larger than the theoretical values ($w(0.9) = 0.235$, $p < 0.001$ for both measures). Thus, empirically, our subjects tend to rely on stronger punishments that are more than sufficient to incentivize the opponents.

In the low accuracy treatment, the retaliation intensity $r(c;0.6) - r(d;0.6)$ is 0.165 in the simple mean of all choices, and 0.144 in the mean of individual-level means. Both measures consistently differ from 0 significantly ($p < 0.001$ for both), which demonstrates that our subjects use signal-contingent information even with the poorer monitoring technology. However, unlike the case of the high accuracy treatment, the retaliation intensity in the low accuracy treatment is much lower than the level implied by the g-TFT equilibria ($w(0.6) = 0.94$) for both measures ($p < 0.001$ for both). Although our subjects do retaliate according to the signals in the low accuracy treatment, the strength of the retaliation is far below the theoretical level, allowing the opponents to defect permanently to pursue larger payoffs. [15]

[Figure 1 HERE]

The retaliation intensities are inconsistent with the values implied by the g-TFT equilibria; further, the deviation is systematic. Since our experimental design is a within-subject design, subject-wise changes of retaliation intensities between the two treatments are observable. Figure 1 displays the histogram of the subject-wise changes of retaliation intensities (i.e., "the retaliation intensity in the high accuracy treatment" minus "the retaliation intensity in the low accuracy treatment"). According to the theoretical prediction, the retaliation intensities should increase drastically from $w(0.9) = 0.235$ in the high accuracy treatment to $w(0.6) = 0.94$ in the low accuracy treatment. However, Figure 1 indicates that none of our subjects follows the theoretical prediction. The differences in retaliation intensities almost cease to exist around -0.715. Almost all of our subjects are either insensitive to signal accuracy and cluster around 0, or show a tendency that is opposite to that predicted by the theory, and cluster around 0.4. A formal statistical test in Table 5 indicates that the retaliation intensity in the high accuracy treatment is larger than the retaliation intensity in the low accuracy treatment ($p < 0.001$ for both the simple mean of all choices and the mean of individual-level means). This is diametrically opposed to the implication of the g-TFT equilibria, wherein people should

[15] There might be concern that the seemingly weaker retaliation intensities observed in this study do not necessarily imply weak retaliation policies of our subjects, since our subjects might use long-term (multi-round) punishments; long-term punishing strategies punish opponents even when observing a good signal during punishing phases, which could lower the retaliation intensity computed here. However, our analysis in Section 8 indicates that only a marginal number of our subjects could adopt long-term punishing strategies, and, hence, the effect of long-term punishments is minimal.

employ weaker retaliating policies in the high accuracy treatment than in the low accuracy treatment.

RESULT 2-a: Our subjects tend to retaliate at a higher level than that implied by the g-TFT equilibria in the high accuracy treatment, while they tend to retaliate less in the low accuracy treatment. Contrary to the implications of the theory, the retaliation intensity is larger when monitoring technology improves.

Remark: One might be concerned that, as the signal becomes less accurate, some subjects simply regard the signal as being uninformative and hence neglect the signal. These subjects would have higher retaliation intensities in the high accuracy treatment than in the low accuracy treatment. If this were the case, all subjects could be divided into two groups; the first group would have zero retaliation intensities in the low accuracy treatment (i.e., higher retaliation intensities in the high accuracy treatment than in the low accuracy one), while the other group would have sufficiently high retaliation intensities in the low accuracy treatment. Our data do *not* support this hypothetical case. Indeed, only a small number of our subjects have higher retaliation intensities in the low accuracy treatment than in the high accuracy treatment. Many subjects have retaliation intensities that are non-negligible but insignificant in the low accuracy treatment. A detailed discussion is provided in Section 8, where we investigate individual strategies.

6.4. Further Results

6.4.1. Reliance on Long Memories

Regarding the behavioral differences of our subjects in relation to signal qualities, it is interesting to examine whether our subjects tend to rely on longer histories of signals (i.e., signals two periods ago). Given the theories regarding review strategies (e.g., Radner, 1986; Matsushima, 2004; Sugaya, 2012), people might rely on signals in longer histories to compensate for the informational disadvantages of poorer monitoring technologies.

To test whether our subjects rely on information in a signal occurring two periods ago, we fit the data with a probabilistic linear regression model, regressing the action choices (a discrete variable that takes the value 1 if our subjects cooperate) on all memory-one histories, which consist of a signal and an action in the previous round, and further, include information of a signal received two periods earlier in the set of explanatory variables, to the extent that non-singularity holds without intercept. [16] The regression coefficients on the signal received two periods earlier capture the additional impact on cooperation probabilities of the signal.

[TABLE 6 HERE]

Table 6 displays the regression results. Contrary to the abovementioned expectations, the regression coefficients on the information received two periods earlier are exclusively significant only in the high accuracy treatment, and one is only marginally significant in the low accuracy treatment. Even in the high accuracy treatment, the size of the coefficients is only marginal, as the maximum impact on cooperative choices is only 13.2%.

RESULT 3-a: There is no evidence that our subjects tend to review longer periods in the low accuracy treatment. Although their actions in the high accuracy treatment partly depend on the information received two periods earlier, the dependencies are marginal at the aggregate level.

6.4.2. Impact of Experiences

Several studies have reported that the frequency of cooperation changes as people experience repeated games (e.g., Dal Bó and Fréchette; 2016). To examine the learning effects of repeated games on overall cooperation rates, we perform a reduced-form linear regression analysis. The results indicate that, despite some learning effects, the size of the

[16] We borrow this approach from Breitmoser (2015).

effects are not remarkably large in our data. The details are reported in Online Appendix E. [17]

7. Estimation of Individual Strategies—Methodology

In this section and the following section, we present direct estimations of individual strategies of our subjects. Given the recent consensus in the literature of experimental repeated games that substantial heterogeneity exists in the strategies employed by subjects (Dal Bó and Fréchette, 2016), we list various strategies and estimate the frequency with which each strategy emerges among our subjects. The primary goal of our exercise is to perform detailed analyses from the viewpoint of individual strategies.

We employ the SFEM methodology developed by Dal Bó and Fréchette (2011). The SFEM is a maximum likelihood estimation (MLE) of a finite mixture model of strategies that subjects use; the model parameters to be estimated are the frequencies of each strategy emerging among subjects, and parameter γ, which controls the stochastic mistakes of action choices or implementation errors, and whose probability of occurrence is $1/(1+\exp(1/\gamma))$. The details of the computation of the likelihood are provided in Appendix B. The underlying assumption is that each subject continues to employ a specific strategy across all repeated games in each treatment. The validity of this method is verified in the Monte Carlo simulations in Fudenberg, Rand, and Dreber (2012). Moreover, we perform simulation exercises to examine the validity of the SFEM that includes a larger set of g-TFT strategies in the strategy list; the results are reported in Online Appendix F.

Because the observed learning effects are not large and are slow to appear, as discussed in Online Appendix E, we employ all the data in the three repeated games for each treatment. As robustness checks, we also perform estimations using only the final two repeated games. The results indicate that there are few changes in the estimates in each treatment; this would not be the case if our subjects systematically changed their strategies across repeated games. The summary of the estimation results is documented in Online Appendix G.

---

[17] See also Section 9.

[TABLE 7 HERE]

Given the difficulty of covering all possible sets of strategies, we include only the strategies that share significant proportions in existing studies on experimental repeated games and their stochastic variations. Table 7 displays the list of strategies in our SFEM. The list includes TFT, TF2T, TF3T, 2TFT, 2TF2T, Grim (trigger strategy),[18] Grim-2, Grim-3, always cooperate (ALL-C), and always defect (ALL-D), all of which are typically listed in the literature of infinitely repeated games in imperfect monitoring (e.g., Fudenberg, Rand, and Dreber, 2012). [19] [20]

Among these, ALL-D is a non-cooperative strategy, while the other strategies (TFT, TF2T, TF3T, 2TFT, 2TF2T, Grim, Grim-2, Grim-3, and ALL-C) are regarded as cooperative strategies that play cooperation at the start of each repeated game and continue to cooperate unless they believe the opponent might switch to defection. TF2T, TF3T, 2TF2T, Grim-2, and Grim-3 are "lenient" strategies (Fudenberg, Rand, and Dreber, 2012) that start punishing only after observing several consecutive occurrences of bad signals. TF2T (TF3T) retaliates once after observing two (three) consecutive bad signals, and 2TF2T retaliates twice after observing two consecutive bad signals; these correspond to a simple form of so-called "review strategies" (lenient strategies with long-term punishments in the proof of the limit folk theorem; see Matsushima, 2004; Sugaya, 2012) Grim-2 (Grim-3) is a lenient variant of Grim strategy, which triggers continuous defection after observing two (three) consecutive deviations from (c, C), that is, the combination of a good signal from the opponent and the subject's own cooperative choice.

---

[18] Here, the definition of the Grim strategy is modified to cover the private monitoring case, in which no common signals are observable. The player starts to continuously choose defection if she observed a bad signal or played defection in the previous round. Note that she could mistakenly play defection before the "trigger" is pulled because implementation errors of action choices are allowed in the SFEM framework.

[19] The literature has often added to the strategy set D-TFT, where defection is played in round 1, followed by TFT in round 2. However, we find no significant frequency of D-TFT in either treatment in our SFEM estimates even if we include D-TFT.

[20] The literature has often added to the strategy set Perfect tit-for-tat (P-TFT), wherein cooperation is played if the both players choose defection in the previous round; otherwise, the action choices are identical to that in TFT. However, we find no significant frequency of P-TFT in either treatment in our SFEM estimates, even if we include P-TFT.

We further include various stochastic strategies, such as g-TFT, g-2TFT, and g-TF2T, in the following manner. Importantly, we add many variants of g-TFT to our SFEM list to identify strategies with various retaliation intensities. We allow the probabilities of cooperation, given a bad signal (i.e., $r(d)$) to take nine distinct values in decrements of 12.5% (i.e., 100%, 87.5%, 75%, 62.5%, 50%, 37.5%, 25%, 12.5%, and 0%). Moreover, we allow the probabilities of cooperation, given a good signal (i.e., $r(c)$), to take nine distinct values in increments of 12.5% (i.e., 100%, 87.5%, 75%, 62.5%, 50%, 37.5%, 25%, 12.5%, and 0%). Here, g-TFT- $r(c)$ - $r(d)$ denotes the g-TFT that plays cooperation after observing a good signal with probability $r(c)$ and after observing a bad signal with probability $r(d)$. [21] We list all possible combinations of $r(c)$ and $r(d)$ in g-TFT in our strategy set, as long as the g-TFT has a non-negative retaliation intensity (i.e., $r(c) \geq r(d)$).

Specifically, we refer to the g-TFT strategies playing cooperation with constant probabilities r, irrespective of the type of signals, as random strategies (denoted by Random-r); they are primitive, signal non-contingent, zero retaliation variants of g-TFT, and include ALL-C and ALL-D as special cases. We regard a g-TFT strategy as non-cooperative if both $r(c)$ and $r(d)$ are no more than 0.5. Otherwise, the g-TFT strategies are cooperative.

We add a family of g-2TFT as strategies that mete out even stronger punishments than TFT. The motivation for this comes from our earlier analysis in Section 6, where we find that our subjects, in aggregate, adopt stronger retaliation intensities than the level implied by the standard theory in the high accuracy treatment. The family of g-2TFT strategies (g-2TFT-r) allows the second retaliations to be stochastic (play cooperation with probability r in the second punishment) as the generous variants of 2TFT. [22]

We also include a family of g-TF2T (g-TF2T-r) as the generous variants of TF2T;

---

[21] For simplicity, we assume that the probability of playing cooperation in round 1 coincides with the choice probability, given a good signal (i.e., $r(c)$).

[22] Unlike g-TFT, g-2TFT does not allow defections until the punishing phases start. Allowing defections outside of punishing phases starts reduces the retaliation intensities, which is contrary to the motivation for employing stronger (multi-round) punishments, rather than punishing in only one round. In addition, the strategies are assumed to play cooperation in round 1, as in TFT; these are multi-round punishing variants of TFT.

theses allow stochastic punishments if two consecutive bad signals occur (play cooperation with probability after observing two consecutive bad signals).[23]

8. Estimation of Individual Strategies—Results

8.1. Cooperative Strategies and Non-Cooperative Strategies

[Table 8 HERE]

[Table 9 HERE]

Table 8 presents the estimates for the frequencies of strategies our subjects follow, and Table 9 displays the aggregated frequencies. In the high accuracy treatment, the share of cooperative strategies, that is, strategies other than ALL-D, Random with a cooperation rate of no more than 0.5, and g-TFT that is less cooperative than Random-0.5 (g-TFT-0.5-$r(d)$, g-TFT-0.375-$r(d)$, g-TFT-0.25-$r(d)$, and g-TFT-0.125-$r(d)$), is 83.9%, while the share of non-cooperative strategies is 16.1%; thus, the share of cooperative strategies exceeds that of non-cooperative strategies. Statistically, the latter share is significantly smaller than the former ($p < 0.001$). Although there are considerable heterogeneities in the strategies that our subjects follow, as Table 8 shows, most of our subjects adopt cooperative strategies in the high accuracy treatment.

On the other hand, in the low accuracy treatment, the share of cooperative strategies is 34.7%, while that of non-cooperative strategies is 65.3%; thus, the share of non-cooperative strategies exceeds that of cooperative strategies. Statistically, the share of non-cooperative strategies in the low accuracy treatment is significantly larger than that in the high accuracy treatment ($p < 0.001$). Among individual strategies, the share of ALL-D is the highest (i.e., 19%). The finding that the share of non-cooperative strategies

[23] Unlike g-TFT, g-TF2T does not allow defections until the punishing phases start; this is because defections outside of punishing phases are contrary to the motivation for employing lenient strategies that allow “giving the benefit of the doubt” to an opponent after the first defection (Fudenberg, Rand, and Druber, 2012). For the same reason, the strategies are assumed to play cooperation in round 1.

is considerable in the low accuracy treatment is consistent with the finding in Section 6. Our SFEM estimates further indicate that more than half of our subjects follow non-cooperative strategies in the low accuracy treatment.

RESULT 1-b: Most of our subjects adopt cooperative strategies in the high accuracy treatment, while in the low accuracy treatment many of them adopt non-cooperative strategies.

Remark: The complexity of the equilibrium cooperative strategy is not the primary impediment to following cooperative strategies in the low accuracy treatment; the equilibrium cooperative strategy in the low accuracy treatment is approximately TFT in our experimental setup; it is quite simple to implement. Rather, our result implies that the poor signal quality in the low accuracy treatment strongly discourages our subjects from adopting cooperative strategies. Fudenberg, Rand, and Dreber (2012) reported in their imperfect public monitoring experiment that frequencies of cooperative strategies drop significantly as the level of noise is increased from 1/16 to 1/8. Our study demonstrates that, even by the more drastic change in signal noise from 1/10 to 4/10 in imperfect private monitoring, the poorer signal quality indeed discourages our subjects from adopting cooperative strategies, even if the experimental parameters are conducive to cooperation.

8.2. Proportion of G-TFT Strategies

We now examine the share of g-TFT. Our SEFM estimates in Tables 8 and 9 indicate that there is a substantial proportion of the family of g-TFT in our imperfect private monitoring. In the high accuracy treatment, g-TFT-1-0.5 has the highest share among the various strategies (17.1%), followed by another g-TFT strategy, g-TFT-0.875-0.25 (10.7%). The total share of the g-TFT family (g-TFT-$r(c)$-$r(d)$), including TFT but excluding signal non-contingent variants of g-TFT (i.e., ALL-C (g-TFT-1-1), ALL-D (g-TFT-0-0), and Random-r (g-TFT-r-r)), is as large as 70.6%. The extended family of g-TFT, which includes the signal non-contingent, primitive variants of g-TFT, has a 76.8% share. Moreover, the further extended family of g-TFT, which includes the family of g-2TFT containing multi-round punishing variants, comprises 77.7% of the strategies. As shown by these large numbers, most of our subjects follow one of the strategies in the g-TFT class.

The share of g-TFT is also substantially large in the low accuracy treatment. The share of the g-TFT family is 55.6%, while the extended family with signal non-contingent variants comprises 94.6% of the strategies; further, the extended family with multi-round punishing variants comprises 96.2% of the strategies. Regardless of the treatment, we find a substantial proportion of our subjects following a strategy from the g-TFT family.

RESULT 4: Our SFEM estimates indicate that the family of g-TFT comprises a substantial proportion of the strategies our subjects follow in both treatments.

This finding indicates that our subjects' decisions on retaliation largely depend on a single occurrence of a bad signal (c.f. lenient strategies). This finding is consistent with our finding in Section 6 that the action choices only marginally depend on the signals that occurred two periods earlier in both treatments. [24]

8.3. Retaliation Intensity

Observing that many of our subjects follow g-TFT, we investigate the proportion of our subjects adopting retaliation intensities that are consistent with g-TFT equilibria. In the high accuracy treatment, g-TFT-1-0.75, g-TFT-0.875-0.625, g-TFT-0.75-0.50, g-TFT-0.625-0.37.5, g-TFT-0.50-0.25, g-TFT-0.375-0.12.5, and g-TFT-0.25-0 have approximately the same retaliation intensity as that implied by the g-TFT equilibria $w(0.9) = 0.235$ in our list. However, our SFEM estimates in Tables 8 and 9 indicate that

[24] The result that our SFEM estimates find many of our subjects playing g-TFT in our imperfect private monitoring is seemingly less consistent with the finding of Fudenberg, Rand, and Dreber (2012), which found only a small proportion of their subjects playing g-TFT in their imperfect public monitoring. However, we are not able to identify the exact factors behind the discrepancy between our results and theirs; this is because our experimental settings different in terms of payoff parameters, discount factor, and signal accuracy. Perhaps most importantly, our setting is a private monitoring environment in which g-TFT plays an important role as an equilibrium strategy. However, as discussed later in Subsection 8.4, we corroborate Fudenberg, Rand, and Dreber (2012) by finding that a certain proportion of our subjects following lenient strategies in the high accuracy treatment.

the joint share of these strategies is only 5.6%, which is not significantly different from 0 ($p = 0.319$); very few of our subjects follow equilibrium g-TFT strategies in the high accuracy treatment. This finding also holds for the low accuracy treatment. In the low accuracy treatment, only TFT has approximately the same retaliation intensity as that implied by the g-TFT equilibria $w(0.6) = 0.94$ in our list. However, the share of TFT in the low accuracy treatment is only 2.7%, which is not significantly different from 0 ($p = 0.221$). These results demonstrate that, despite many of our subjects following one of the g-TFT strategies, almost none follows the retaliation intensities implied in the g-TFT equilibria in both treatments.

We further address how they tend to deviate from the theoretical prediction in terms of proportions of strategies. Our SFEM estimates in Tables 8 and 9 indicate that the group of stronger retaliation variants of g-TFT, that is, g-TFT strategies with retaliation intensities of more than 0.25, comprising g-TFT-1-0.625/0.5/0.375/0.25/0.125/0, g-TFT-0.875-0.5/0.375/0.25/0.125/0, g-TFT-0.75-0.375/0.25/0.125/0, g-TFT-0.625-0.25/0.125/0, g-TFT-0.5-0.125/0, g-TFT-0.375-0, and g-2TFT-0.875/0.75/0.625/0.5/0.375/0.25/0.125/0, jointly comprises 54.4% of the strategies in the high accuracy treatment. Since slightly more than 70% of our subjects follow g-TFT, the results imply that roughly three-fourths of them retaliate more strongly than the g-TFT equilibria require. Indeed, the share of the weaker retaliation variants of g-TFT, that is, g-TFT-1-0.875, g-TFT-0.875-0.75, g-TFT-0.75-0.625, g-TFT-0.625-0.5, g-TFT-0.5-0.375, g-TFT-0.375-0.25, g-TFT-0.25-0.125, g-TFT-0.125-0, ALL-C, ALL-D, and Random-r, merely reaches 17.7%; statistically, thus is significantly smaller than the share of the group of stronger retaliation variants ($p < 0.001$).

On the other hand, in the low accuracy treatment, our SFEM estimates indicate that the group of weaker retaliation variants of g-TFT, that is, g-TFT strategies other than TFT, ALL-C, ALL-D, and Random-r, jointly comprise 93.5%. Strong retaliation variants of g-TFT comprise less than 0.1% of the strategies; statistically, is this is significantly smaller than the share of the weaker retaliation variants ($p < 0.001$).

In Section 6, we found that the mean retaliation intensities deviate from the values implied by the g-TFT equilibria at the aggregate level. We examine whether this finding holds even if we restrict our attention to the behavior of g-TFT players, rather than considering the aggregate behavior for all the strategies. We compute the mean retaliation intensities conditional on the g-TFT strategies (including ALL-C, ALL-D, and Random-r). The conditional mean retaliation intensity in the high accuracy treatment is 0.426 (s.e. 0.033); this is significantly larger than the value predicted by the g-TFT equilibria (0.235, $p < 0.001$). In the low accuracy treatment, the mean retaliation intensity is 0.148 (s.e. 0.025) ; this is significantly smaller than the value implied by the g-TFT equilibria (0.94, $p < 0.001$). The conditional mean of retaliation intensities in the high accuracy treatment is significantly larger than that in the low accuracy treatment ($p < 0.001$). Hence. even if we consider only g-TFT players, the behavior systematically deviates from the theoretical predictions; this result is similar to the findings in Section 6.

RESULT 2-b: Our SFEM estimates indicate that, in both treatments, only a small number of our subjects follow the retaliation intensities implied by the g-TFT equilibria. In the high accuracy treatment, the share of stronger retaliation variants of g-TFT outweighs that of weaker retaliation variants. In the low accuracy treatment, the share of weaker retaliation variants of g-TFT outweighs that of stronger retaliation variants. The mean retaliation intensity among g-TFT players is larger than the value implied by the g-TFT equilibria in the high accuracy treatment, and is smaller than the value implied by the g-TFT equilibria in the low accuracy treatment. The mean retaliation intensity of the g-TFT strategies in the high accuracy treatment is larger than that in the low accuracy treatment; this is contrary to the theoretical implications of the g-TFT equilibria.

Remark: One might be concerned that some subjects simply regard the signal as being uninformative and neglect the signal as the signal accuracy becomes deteriorated. If this occurred, all subjects in the low accuracy treatment could be divided into two groups; one group would have zero retaliation intensities, while the other group would have sufficiently high retaliation intensities. To clarify whether our experimental results support this idea, we first compute the fraction of our subjects who follow signal non-contingent, zero retaliation strategies (i.e., ALL-C, ALL-D, and Random strategies) in the low accuracy treatment. For each subject, we compute a Bayesian posterior for employing the signal non-contingent strategies and choose subjects whose posterior is more than

0.95. [25] The estimates indicate that only a small number of subjects (i.e., 25 out of 108) follow the signal non-contingent, zero retaliation strategies. For the other 83 subjects, the mean of retaliation intensities (individual-level) is 0.1883; this is still far below the value implied by the g-TFT equilibria. Indeed, only a few subjects in the subsample have strong retaliation intensities in the low accuracy treatment. Only three subjects have strong retaliation intensities that exceed 0.9 (i.e., the value close to that implied by the g-TFT equilibria); six subjects have retaliation intensities of more than 0.5. Second, we compute the mean of the difference of retaliation intensities (individual-level) in this 83-subject subsample (histogram is displayed in Figure 1). We obtain a value of 0.1900, which is close to the original value presented in Table 5; this implies that our finding is robust to the existence of the signal non-contingent strategies.

8.4. Further Results

8.4.1. Long-Term Punishment

We now focus on long-term punishing strategies. In Section 6, we found that the aggregate level of retaliation intensity is smaller than the level implied by the g-TFT equilibria in the low accuracy treatment. We address the concern that seemingly weak retaliation intensity might spuriously arise when some subjects employ long-term punishing strategies. Tables 8 and 9 indicate that this is not the case. The share of strategies with long-term punishments, that is, a family of 2TFT, 2TF2T, Grim, Grim-2, and Grim-3, in the low accuracy treatment is less than 3%, which is statistically insignificant ($p = 0.435$); the effect of strategies with long-term punishment is minimal. This is also true in the high accuracy case; the joint share of the long-term punishing strategies is 6.6%, which is statistically insignificant ($p = 0.203$).

RESULT 2-c: The share of strategies with long-term punishments that could spuriously reduce the observed retaliation intensities is small in both treatments.

---

[25] Other threshold values, such as 0.90 or 0.80, do not alter our conclusion significantly.

8.4.2. Lenience

We focus on the share of lenient strategies. As seen in Section 6, our subjects do not tend to rely on longer histories in the low accuracy treatment. This suggests that, contrary to speculations in the field of review strategies, the share of lenient strategies might not be larger when the monitoring technologies are poorer (Radner, 1986; Matsushima, 2004; Sugaya, 2012). Tables 8 and 9 indicate that, in the low accuracy treatment, the individual shares of none of the lenient (review) strategies, that is, the family of g-TF2T including TF2T, TF3T, 2TF2T, Grim-2, and Grim-3, are significantly different from 0. Taken together, they comprise only 3.8% of the strategies chosen; this share is not significantly different from 0 ($p = 0.359$). The lenient strategies do not comprise a remarkable proportion in the low accuracy treatment. In contrast, the share of lenient strategies in the high accuracy treatment, rises to 21.4%, which is significantly different from 0 ($p = 0.023$); further, it is marginally significantly larger than that of the low accuracy treatment ($p = 0.055$). Thus, there is no tendency to adopt more lenient strategies when monitoring technologies are poorer.

RESULTS 3-b: Our SFEM estimates provide no evidence of a larger share of lenient strategies in the low accuracy treatment. Instead, the share is negligible only in the low accuracy treatment, while it is approximately 20% in the high accuracy treatment.

9. Alternative Experimental Design[26]

9.1 Aims of the Design and Summary of the Results

Subjects' choices of strategies could be affected by various factors, including slight changes in experimental designs. It is meaningful to examine whether an alternative experimental design could bring any change of subjects' behavior as a robustness check

---

[26] We are grateful to an anonymous referee for the suggestions about additional experiments.

of our experimental findings.

In the main experiments, we adopted a within-subject design in which each subject is assigned to both the experimental treatments, that is, high and low accuracy treatments, which enables us to observe subject-wise changes of behavior. However, due to a time constraint, each subject experienced only three repeated games per treatment. Thus, one might argue that the learning of the subjects was still in progress, and other strategies could have emerged if the experiment had continued. [27] On the other hand, if subjects experience only a single treatment many times, they could get bored and become reluctant to engage seriously in experimental tasks. In the worst case, subjects could play "inertia". To address both these drawbacks, we conducted additional experiments with between-subjects design, and report the results in this section.

Further, in the theoretical study of repeated games, there are two interpretations on infinite repetition of games; one interpretation assumes that the games should be repeated physically infinitely as the way in which canonical theories are developed (of course, which does not occur strictly in reality). The other interpretation does not require ultimately infinite continuation of games in reality, but regards that players perceive finite strategic interactions but repeating many times as infinite repetitions (e.g., Osborne and Rubinstein, 1994, Chapter 8). Our main experimental design adopts the latter interpretation, and we honestly announced a quite large value (i.e., 98) as the maximum number of rounds for each repeated game with random termination.

However, it would be helpful to attempt alternative experimental design based on the former interpretation wherein we do not announce the maximum number of rounds. However, since we cannot continue repeated games ultimately infinitely, we can carry out the experiments only in case that no objection or question is raised by subjects about the absence of an announcement regarding the maximum number of rounds (hence, the maximum termination time of the experiment) the subjects would undergo. [28] Fortunately, we encountered no objection or question about it from the subjects in the

[27] For the learning effects in the main experiments, see Subsection 6.5.2.

[28] In a pilot experiment during the previous experiments a subject, perhaps worried about the termination time of the experiment, had asked an experimenter whether there is an upper limit to the number of games. This is a natural and primary concern of subjects. Being wary of facing a similar question in other sessions, and being considered non-transparent if we did not answer it, we announced the same maximum number of rounds in every session to ensure that the subjects were treated equally during the previous experiments.

new experiments.

In the new experiments, we change the parameter in the low accuracy to $p = 0.7$; this is done so that we might observe inertia players than the case when $p = 0.6$. It is worth noting that because of the change of the low accuracy treatment parameter, the theoretical value of retaliation intensity implied by g-TFT equilibria $w(0.7)$ in the low accuracy treatment is 0.470.

To summarize the experimental results, we find that they are mostly consistent with our previous results; our SFEM estimates indicate that a class of g-TFT strategies account for a substantial proportion of strategies in both treatments. The mean of the observed retaliation intensities does not increase as the monitoring accuracy becomes poorer. The novel result in the new experiments was the emergence of a certain proportion of subjects playing Lenience, Long-term punishment, Grim, and a hybrid of Lenience and g-TFT strategies (i.e., g-TF2T). These strategies were either nonexistent or constituted small proportions in the previous experiments. The results observed in the new experiments show that, although g-TFT shares substantial portions, subjects' choices of strategies are affected more or less by the way of designing experiments, even though the difference of designs is relatively slight one.

9.2 Details of the Experiments

We conducted four sessions at the Center for Advanced Research for Finance, University of Tokyo in August 2018. We recruited 112 subjects from a subject pool consisting of undergraduate and graduate students from various fields. Fifty-six subjects were assigned to the high accuracy treatment, and the other fifty-six were assigned to the low accuracy treatment.

[Table 10 HERE]

Each subject plays five repeated games in the assigned treatment. Unlike the previous experiments, the number of rounds in the repeated games are determined in advance. Using the same continuation probability as the previous experiments (i.e., $\delta = 29/30$)

we randomly generate two sequences of rounds for five repeated games, and we use each sequence once in each treatment. This procedure matches the sequences of repeated games across two treatments, thereby erasing possible heterogeneities in the learning effects owing to randomness of the number of rounds experienced across treatments. Table 10 presents the full sequences of repeated games in the experiments. The payoff parameter is the same as in the previous one (presented in Table 1). [29] [30]

[Table 11 HERE]

[Table 12 HERE]

[Table 13 HERE]

Table 11 displays the descriptive summary of the data; Table 12 presents the round 1 cooperation rates and signal-contingent cooperation rates; and Table 13 presents retaliation intensities. As in the previous analysis performed in Section 6, we primarily focus on the means of individual-level means instead of simple means of action choices to avoid over-representations of particular types of subjects.

In the previous experiments, the values for the mean of the retaliation intensities were found to be slightly larger than the value implied by the g-TFT equilibria in the high accuracy treatment; here, the mean of the retaliation intensities in the experiment is 0.281. Statistically, this value is not significantly different from the value implied by the theory (0.235, $p = 0.127$). On the other hand, in the low accuracy treatment, as in the previous experiments, we find that the mean of the retaliation intensities is 0.249, which is significantly smaller than the value implied by the g-TFT equilibria (0.470, $p < 0.001$). The direct comparison of the two retaliation intensities indicates that the retaliation intensity in the low accuracy treatment is not significantly larger than that in the high

[29] Although the payoff matrix is identical, the conversion rate from the earned points in the experiment to the monetary reward is changed from 0.6 to 0.9, since the students' earnings in this area have generally risen in recent years.

[30] The experimental instructions for the new experiment are provided in Appendix I.

accuracy treatment ($p = 0.458$). [31]

As for the learning effects, we performed an analysis similar to that reported in Subsection 6.4.2 and Online Appendix E; the fixed-effect model analysis was performed by regressing the cooperation rates on dummy variables for repeated games. Table 14 reports the regression results. The dummy variable *RG2*/*RG3/RG4/RG5* in the regressor takes a value of 1 if the choice is made in the second/third/forth/fifth repeated game.

[Table 14 HERE]

Although the difference in the coefficient on the regressor *RG5* from that on the regressor *RG3* in the high accuracy treatment is statistically (but marginally) significant ($p = 0.090$), the coefficient on the regressor *RG4* does not ($p = 0.159$). This is also true in the low accuracy treatment in which the coefficient on the regressor *RG5* differs (but marginally) significantly from that on the regressor *RG3* ($p = 0.066$), but does not differ from that on the regressor *RG4* ($p = 0.553$). These results suggest that the frequencies of cooperation do not change drastically across the fourth and fifth repeated games in both treatments. Observing this, we employ the data generated in the final two repeated games for the SFEM analysis. [32]

The list of strategies employed in the SFEM performed here is identical to that in the previous SFEM, which is shown in Table 7. However, the increment in choice probabilities in the stochastic strategies in the estimations is now 25% instead of 12.5%; this is because the sample size in each treatment of the experiments is reduced to almost half of that in the previous experiments owing to the between-subjects experimental design. Then the probabilities of cooperation in g-TFT strategies, g-2TFT strategies, and g-TF2T strategies take five distinct values in decrements of 25%, that is, 100%, 75%, 50%, 25%, and 0%.

[31] There is no qualitative change in these results even if we use only the data of the final two repeated games.

[32] Although the cooperation rates differ significantly across the entire five repeated games, perhaps due to the small absolute size of the learning effects, the estimation results in the SFEM differ only slightly even if we use all the data. Conservatively, we report the results based on the final two repeated games.

[Table 15 HERE]

[Table 16 HERE]

Table 15 presents the estimates for the frequencies of various strategies, and Table 16 displays the aggregated frequencies. In the low accuracy treatment, g-TFT-0.75-0.5 has the highest individual share (17.7%), and the joint share of the strategies classified as g-TFT is 72.7% (s.e. 15.3%); statistically, this is significantly larger than zero ($p < 0.001$). In the high accuracy treatment, although the largest individual frequency is that of g-TF2T-0.5 (27.3%), the family of g-TFT has a large frequency. They constitute 51.2% (s.e. 10.1%) of the strategies; statistically, this is significantly larger than zero ($p < 0.001$).

The mean retaliation intensity among g-TFT players in the low accuracy treatment is 0.235 (s.e. 0.057) ; statistically, this is significantly smaller than the theoretical value $w(0.7) = 0.470$ ($p < 0.001$). In the high accuracy treatment, the point estimate is 0.251 (s.e. 0.063), which is only slightly larger than the value implied by the theory $w(0.9) = 0.235$, although the difference is insignificant ($p = 0.798$). The direct comparison of the retaliation intensities across the two treatments indicates that the retaliation intensity in the high accuracy treatment is not statistically smaller than that in the low accuracy treatment ($p = 0.849$).

We find somewhat larger frequencies of long-term punishment strategies in the experiments. The joint share of the long-term punishment strategies is 16.0% ($p = 0.038$, s.e. 7.7%) in the high accuracy treatment, and 20.2% ($p = 0.050$, s.e. 10.3%) in the low accuracy treatment. As evaluated from the point estimates, Grim-3 appears to account for a major share. In the low accuracy treatment, the individual share of the Grim-3 strategy is 14.6% and the value is significantly different from zero ($p = 0.010$). In the high accuracy treatment, the point estimate of the share of Grim-3 strategy is 10.9%; however, statistically, the value is not significantly different from zero owing to a large standard error ($p = 0.171$).

Further, we find larger frequencies of lenient strategies in the experiments; however, as in the previous results, their share does not increase as the monitoring becomes poorer. In the current experiments, the frequency of lenient strategies in the high accuracy treatments is 47.0% ($p < 0.001$, s.e. 10.3%), and that in the low accuracy treatment is 24.6% ($p = 0.043$, s.e. 12.1%). The major share in the latter is again that of Grim-3. In the former, the g-TF2T-0.5 strategy has a large share (27.3%). The point estimates indicate that the share of lenient strategies does not increase as monitoring becomes poorer, though the difference is insignificant ($p = 0.164$) owing to a large standard error.

10. Conclusion

This study experimentally examines collusion in a repeated prisoner's dilemma game, with random termination, in which monitoring is imperfect and private. Each player obtains information about the opponent's action choice through a signal instead of a direct observation, and the signal the opponent observes is not observable by the player. The continuation probability in the experiments is large enough to allow the subjects to collude even if the monitoring technology is poor. Our study is the first experimental attempt to investigate imperfect private monitoring.

Our experimental results indicate that a significant proportion of our subjects employed g-TFT strategies, which are straightforward stochastic extensions of the TFT strategy. We depart significantly from the experimental literature by considering g-TFT strategies, which have attracted less attention in the empirical literature despite their theoretical and practical importance. Our finding that a significant proportion of our subjects follow g-TFT strategies reveals its empirical importance. Our estimation results indicate that a large proportion of the subjects follow g-TFT strategies, rather than grim-trigger, lenience, long-term punishment strategies, and their stochastic variants, that frequently appear in existing experimental studies of repeated games.

Although many subjects follow g-TFT strategies, their retaliating policies deviate systematically from the predictions of the g-TFT equilibria. Our subjects retaliate more in the high accuracy treatment than in the low accuracy treatment; this is contrary to the theoretical predictions. The subjects retaliate more in the high accuracy treatment, while in the low accuracy treatment, they retaliate less than what the standard g-TFT equilibria predict. Hence, the expected payoff to a subject from cooperation tends to be much greater than the expected payoff from defection when monitoring is accurate; however, the expected payoff from cooperation tends to be much less than the expected payoff from defection when monitoring is inaccurate.

These experimental findings indicate that subjects fail to improve their welfare by effectively utilizing monitoring technology, as predicted by the standard theory. Rather, this systematic deviation from the standard theory might suggest to economists a new issue related to incentives that encompasses the motivations for retaliations beyond just maximizing pay-off. [33]

Further, it is meaningful to attempt various experimental designs and examine subjects' choice of strategies in a more systematic manner. From this study and the previous experimental ones concerning imperfect monitoring, such as Fudenberg, Rand, and Dreber (2012), we learnt that g-TFT and lenient strategies are remarkable in implicit collusion. However, grim-trigger and long-term punishing strategies are not much even though they are prominent not only in theory, but also in real cartels (e.g., Igami and Sugaya, 2018). Hence, as future researches it would be fruitful to investigate various potential factors such as pre-play communication and information transmission during the play, that could affect players' choices of strategies experimentally; these are expected to make experimental environments resemble reality. [34]

References

Abreu, D. (1988): "On the Theory of Finitely Repeated Games with Discounting," *Econometrica* 56, 383–396.

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* 58, 1041–1063.

Aoyagi, M., V. Bhaskar, and G. Fréchette (2019): "The Impact of Monitoring in Infinitely Repeated Games: Perfect, Public, and Private," *American Economic Journal:*

[33] The companion paper Kayaba, Matsushima, and Toyama (2018) demonstrates a behavioral model of an infinitely repeated prisoners' dilemma.

[34] Some of us are preparing follow-up experimental studies that incorporate these factors.

Microeconomics 11(1), 1-43.

Aoyagi, M. and G. Fréchette (2009): "Collusion as Public Monitoring Becomes Noisy: Experimental Evidence," *Journal of Economic Theory* 144, 1135-1165.

Axelrod, R. (1984): *The Evolution of Cooperation*, Basic Books: New York.

Binmore. K. (1994): *Game Theory and the Social Contract, Vol. 1: Playing Fair*, MIT Press: Cambridge.

Breitmoser, Y. (2015): "Cooperation, but No Reciprocity: Individual Strategies in the Repeated Prisoner's Dilemma," *American Economic Review* 105(9), 2882–2910.

Dal Bó, P. (2005): "Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games," *American Economic Review* 95, 1591–1604.

Dal Bó, P. and G. Fréchette (2011): "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *American Economic Review* 101(1), 411–429.

Dal Bó, P. and G. Fréchette (2016): "On the Determinants of Cooperation in Infinitely Repeated Games: A Survey," *Journal of Economic Literature*, forthcoming.

Ely, J. and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory* 102, 84–105.

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics* 10(2), 171–178.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica* 62, 997–1040.

Fudenberg, D. and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting and with Incomplete Information," *Econometrica* 54, 533–554.

Fudenberg, D., D.G. Rand, and A. Dreber (2012): "Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World," *American Economic Review* 102(2), 720–749.

Green, E. and R. Porter (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica* 52, 87–100.

Igami, M. and T. Sugaya (2018): "Measuring the Incentive to Collude: The Vitamin Cartels, 1990-1999," Working Paper.

Kayaba, Y., H. Matsushima, and T. Toyama (2018): "Psychological Aspects of Monitoring Accuracy in Repeated Prisoners' Dilemma," Discussion Paper, CARF-F-432, University of Tokyo.

Mailath, J. and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run*

Relationships, Oxford University Press.

Matsushima, H. (2004): "Repeated Games with Private Monitoring: Two Players," *Econometrica* 72, 823–852.

Matsushima, H. (2013): "Interlinkage and Generous Tit-For-Tat Strategy," *Japanese Economic Review* 65, 116–121.

Matsushima, H., T. Tanaka, and T. Toyama (2013): "Behavioral Approach to Repeated Games with Private Monitoring," DP CIRJE-F-879, University of Tokyo.

Matsushima, H. and T. Toyama (2011): "Monitoring Accuracy and Retaliation in Infinitely Repeated Games with Imperfect Private Monitoring: Theory and Experiments," DP CIRJE-F-795, University of Tokyo.

Matsushima, H., T. Toyama, and N. Yagi (2007): "Repeated Games with Imperfect Monitoring: An Experimental Approach," University of Tokyo.

Molander, P. (1985): The Optimal Level of Generosity in a Selfish Uncertain Environment, *Journal of Conflict Resolution* 29, 611-618.

Nowak, M. and K. Sigmund (1992): "Tit-For-Tat in Heterogeneous Populations," *Nature* 355, 250–253.

Osborne, M. and A. Rubinstein (1994): *A Course in Game Theory*, The MIT Press: Cambridge.

Piccione, M. (2002): "The Repeated Prisoners' Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory* 102, 70–83.

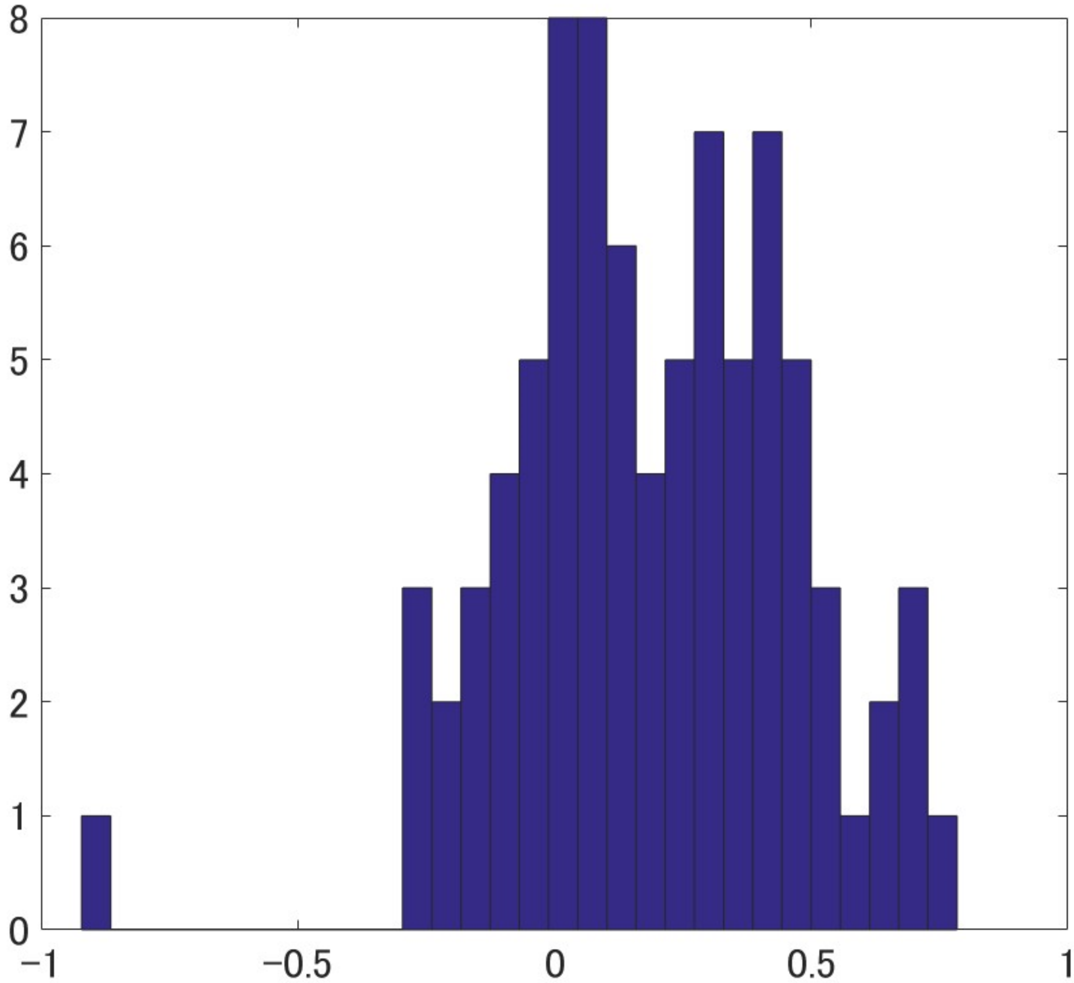
Radner, R. (1986): "Repeated Partnership Games with Imperfect Monitoring and No Discounting," *Review of Economic Studies* 53, 43–57.

Sugaya, T. (2012): *The Folk Theorem in Repeated Games with Private Monitoring*, Ph. D. dissertation.

Takahashi, S. (2010): "Community Enforcement when Players Observe Partners' Past Play," *Journal of Economic Theory* 145, 42–64.

Figure 1:

Histogram of Subject-wise Changes of Retaliation Intensities



X-axis: RI in High Accuracy minus RI in Low Accuracy
Twenty-six subjects who are considered to be employing signal non-contingent strategies in the low accuracy treatment are excluded (see discussion in Section 8.3).

Table 1:
Prisoner's Dilemma with Symmetry and Additive Separability

	C	D
C	60 60	5 70
D	70 5	15 15

Table 2:
Features of Experimental Design

	Number of subjects	Treatment order (sequence of game lengths)
October 5, 2006 (10:30–12:30)	28	0.6 (24, 40, 25), 0.9 (28, 33, 14)
October 5, 2006 (14:30–16:30)	24	0.6 (20, 23, 37), 0.9 (34, 34, 19)
October 6, 2006 (10:30–12:30)	28	0.9 (38, 21, 25), 0.6 (25, 28, 29)
October 6, 2006 (14:30–16:30)	28	0.9 (25, 35, 23), 0.6 (36, 30, 21)

Table 3:
Decisions and Signal

	$p = 0.9$			$p = 0.6$		
	N	Mean	St. Dev.	N	Mean	St. Dev.
Cooperative choice	8,864	0.672	0.469	9,144	0.355	0.479
Good signal	8,864	0.637	0.481	9,144	0.483	0.500

Table 4:
Means of Cooperative Action Choice

	$p = 0.9$	$p = 0.6$	p-values
$q(p)$ (round 1)	0.781 (0.042)	0.438 (0.035)	< 0.001
$r(c; p)$	0.852 (0.033)	0.437 (0.018)	< 0.001
Individual-level means	0.799 (0.032)	0.423 (0.027)	< 0.001+
$r(d; p)$	0.344 (0.026)	0.272 (0.030)	< 0.001
Individual-level means	0.448 (0.026)	0.279 (0.032)	< 0.001+

Notes: The standard errors (shown in parentheses) are block-bootstrapped (individual and repeated game level) with 5,000 repetitions, which is used to calculate p-values. The null hypothesis is that the values are identical across the two treatments.
+ The Wilcoxon matched-pair test within individual rejects the null hypothesis that the values are identical across the two treatments ($p < 0.001$ for each).

Table 5:
Retaliation Intensities

	Mean	*S.E.*	*p-value*
$r(c;0.9)-r(d;0.9)$	0.508	0.028	< 0.001+
Individual-level means	0.352	0.028	< 0.001+
$r(c;0.6)-r(d;0.6)$	0.165	0.025	< 0.001+
Individual-level means	0.144	0.023	< 0.001+
$(r(c;0.9)-r(d;0.9))-(r(c;0.6)-r(d;0.6))$	0.344	0.031	< 0.001
Individual-level means	0.208	0.029	< 0.001

Notes: The standard errors are block-bootstrapped (subject and repeated game level) with 5,000 repetitions, which is used to calculate p-values.
+ The hypothesis tests for the comparison to the value implied by the standard theory ($w(p)$), which is 0.235 in the high accuracy treatment and 0.94 in the low accuracy treatment. The null hypothesis is that the mean is identical to the implied value.

Table 6:
Reduced-form Regression Results on Signal Contingency

	p = 0.9	**p = 0.6**
cC	0.887*** (0.017)	0.766*** (0.035)
cD	0.502*** (0.036)	0.255*** (0.032)
dC	0.506*** (0.033)	0.582*** (0.044)
dD	0.106*** (0.028)	0.110*** (0.015)
cCc	0.062*** (0.015)	0.037 (0.027)
cDc	0.017 (0.047)	-0.055* (0.029)
dCc	0.125*** (0.035)	-0.045 (0.042)
dDc	0.132*** (0.032)	0.022 (0.013)
Observations	8,216	8,496
R2	0.816	0.531

Notes: xYz in the regressors denote that the player takes action Y, observes signal x about the opponent's choice, and observed signal z in the previous round. Similarly, xY denotes that the player takes action Y and observes signal x about the opponent's choice. Cluster-robust (individual-level) standard errors in parenthesis. $*p<0.1$, $**p<0.05$, $***p<0.01$.

Table 7:
Strategy Set in our SFEM

Strategy	Description
ALL-C	Always cooperate
TFT	Tit-for-tat
g-TFT- $r(c)$ **-** $r(d)$	Generous tit-for-tat (cooperate if a good signal occurs with probability $r(c)$; forgive in the event of a bad signal and cooperate with probability $r(d)$)
ALL-D	Always defect
TF2T	Tit-for-two-tat (retaliate if bad signals occur in both of the last two rounds)
g-TF2T- r	Generous tit-for-two-tat (play cooperation stochastically with probability r even after observing two consecutive bad signals)
TF3T	Tit-for-three-tat (retaliate if a bad signal occurs in all of the last three rounds)
2TFT	Two tit-for-tat (retaliate twice consecutively if a bad signal occurs)
g-2TFT- r	Generous two tit-for-tat (certainly retaliate if a bad signal occurs, but forgive and cooperate with probability r in the next round if a good signal occurs (second punishment))
2TF2T	Two tit-for-two-tat (retaliate twice consecutively if a bad signal occurs in both of the last two rounds)
Grim	Cooperate until the player chooses defection or observes a bad signal, and then play defection forever
Grim-2	Cooperate until the case of twice in a row occurs, in which the player chooses defection or observes a bad signal, and then play defection forever
Grim-3	Cooperate until the case of three times in a row occurs, in which the player chooses defection or observes a bad signal, and then play defection forever
Random- r	Cooperate with probability r irrespective of signals

Table 8:
Maximum Likelihood Estimates of Individual Strategies

	p = 0.9	S.E.	p = 0.6	S.E.
ALL-C (g-TFT-1-1)	0	(0.029)	0.037**	(0.019)
g-TFT-1-0.875	0	(0.009)	0	(0)
g-TFT-1-0.75	0	(0)	0	(0.002)
g-TFT-1-0.625	0	(0.040)	0	(0)
g-TFT-1-0.5	0.171***	(0.063)	0	(0)
g-TFT-1-0.375	0	(0)	0	(0)
g-TFT-1-0.25	0.041	(0.031)	0	(0)
g-TFT-1-0.125	0	(0.007)	0.010	(0.014)
TFT (g-TFT-1-0)	0.026	(0.016)	0.027	(0.022)
Random-0.875 (g-TFT-0.875-0.875)	0.002	(0.014)	0	(0)
g-TFT-0.875-0.75	0.033	(0.024)	0.027	(0.017)
g-TFT-0.875-0.625	0.022	(0.033)	0	(0.015)
g-TFT-0.875-0.5	0.042	(0.041)	0.004	(0.015)
g-TFT-0.875-0.375	0.055	(0.048)	0	(0)
g-TFT-0.875-0.25	0.107***	(0.041)	0	(0)
g-TFT-0.875-0.125	0	(0.006)	0.010	(0.011)
g-TFT-0.875-0	0	(0)	0.010	(0.009)
Random-0.75 (g-TFT-0.75-0.75)	0	(0)	0	(0)
g-TFT-0.75-0.625	0	(0.001)	0.038	(0.024)
g-TFT-0.75-0.5	0	(0.017)	0.017	(0.022)
g-TFT-0.75-0.375	0.003	(0.023)	0.026	(0.027)
g-TFT-0.75-0.25	0	(0)	0	(0.004)
g-TFT-0.75-0.125	0.011	(0.015)	0	(0)
g-TFT-0.75-0	0.009	(0.008)	0	(0)
Random-0.625 (g-TFT-0.625-0.625)	0.011	(0.010)	0.014	(0.018)
g-TFT-0.625-0.5	0	(0.005)	0	(0.012)
g-TFT-0.625-0.375	0.023	(0.029)	0.089*	(0.046)
g-TFT-0.625-0.25	0	(0)	0	(0.015)
g-TFT-0.625-0.125	0.026	(0.021)	0	(0)
g-TFT-0.625-0	0	(0.010)	0	(0)
Random-0.5 (g-TFT-0.5-0.5)	0	(0.013)	0.023	(0.024)
g-TFT-0.5-0.375	0.053	(0.034)	0.050	(0.045)
g-TFT-0.5-0.25	0.003	(0.011)	0.038	(0.029)
g-TFT-0.5-0.125	0	(0)	0	(0.006)
g-TFT-0.5-0	0.019	(0.014)	0	(0)
Random-0.375 (g-TFT-0.375-0.375)	0.021	(0.019)	0.044	(0.045)
g-TFT-0.375-0.25	0	(0.006)	0.090	(0.060)
g-TFT-0.375-0.125	0	(0)	0	(0.012)
g-TFT-0.375-0	0	(0.003)	0.014	(0.013)
Random-0.25 (g-TFT-0.25-0.25)	0	(0)	0.063	(0.053)
g-TFT-0.25-0.125	0	(0)	0.069*	(0.041)
g-TFT-0.25-0	0.008	(0.014)	0.001	(0.011)
Random-0.125 (g-TFT-0.125-0.125)	0	(0.001)	0.033	(0.028)
g-TFT-0.125-0	0.029	(0.021)	0.037	(0.022)
ALL-D (g-TFT-0-0)	0.028	(0.017)	0.190***	(0.035)
g-TF2T-0.875	0.074*	(0.040)	0	(0)
g-TF2T-0.75	0	(0.021)	0	(0)
g-TF2T-0.625	0	(0.010)	0	(0.003)
g-TF2T-0.50	0.022	(0.028)	0	(0)
g-TF2T-0.375	0	(0.020)	0	(0)
g-TF2T-0.25	0	(0.024)	0.009	(0.009)

g-TF2T-0.125	0	(0.001)	0	(0)
TF2T (g-TF2T-0)	0	(0.004)	0	(0)
TF3T	0.070	(0.045)	0	(0.004)
g-2TFT-0.875	0	(0)	0	(0)
g-2TFT-0.75	0.009	(0.016)	0	(0)
g-2TFT-0.625	0	(0.024)	0	(0)
g-2TFT-0.50	0	(0.005)	0	(0)
g-2TFT-0.375	0	(0)	0	(0)
g-2TFT-0.25	0	(0)	0	(0)
g-2TFT-0.125	0	(0)	0	(0)
2TFT (g-2TFT-0)	0	(0)	0	(0)
2TF2T	0	(0.013)	0.009	(0.010)
Grim	0.009	(0.008)	0	(0)
Grim-2	0.013	(0.016)	0	(0.007)
Grim-3	0.035	(0.025)	0.020	(0.024)
Gamma	0.280***	(0.035)	0.247***	(0.059)

Notes: The standard errors (shown in parentheses) are cluster-bootstrapped (individual level) with 100 repetitions, which is used to calculate p-values. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table 9:
Frequency of Cooperative Strategies, g-TFT, Retaliation Intensities (RI)

		p = 0.9		p = 0.6
Cooperative strategies		83.9%		34.7%
Non-cooperative strategies		16.1%		65.3%
Family of g-TFT		76.8%		96.2%
Excluding signal non-contingent, zero RI strategies (ALL-C, ALL-D, Random-r)		70.6%		55.6%
Including multi-round punishing strategies (g-2TFT)		77.7%		96.2%
Equilibrium RI	(RI = 0.25)	5.6%	(RI = 1)	2.7%
Stronger RI	(RI > 0.25)	54.4%	(RI > 1)	0.0%
Weaker RI (positive RI only)	(0 < RI < 0.25)	11.4%	(0 < RI < 1)	53.0%
Including zero RI strategies	(RI < 0.25)	17.7%	(RI < 1)	93.5%
Mean RI conditional on g-TFT		0.426		0.148
Median RI conditional on g-TFT		0.5		0.125
Lenient strategies		21.4%		3.7%

Table 10:
Features of Experimental Design

	Number of subjects	Treatment (sequence of game lengths)
October 2, 2018 (10:30–12:30)	28	0.9 (11, 34, 48, 21, 33)
October 3, 2018 (10:30–12:30)	28	0.9 (42, 11, 30, 46, 24)
October 3, 2018 (14:00–16:00)	28	0.7 (11, 34, 48, 21, 33)
October 4, 2018 (10:30–12:30)	28	0.7 (42, 11, 30, 46, 24)

Table 11:
Decisions and Signal

	$p = 0.9$			$p = 0.7$		
	N	Mean	St. Dev.	N	Mean	St. Dev.
Cooperative choice	8,400	0.847	0.360	8,400	0.585	0.493
Good signal	8,400	0.777	0.417	8,400	0.534	0.499

Table 12:
Means of Cooperative Action Choice

	$p = 0.9$	$p = 0.7$	p-values
$q(p)$ (round 1)	0.946 (0.026)	0.780 (0.047)	< 0.001
$r(c; p)$	0.928 (0.013)	0.719 (0.038)	< 0.001
Individual-level means	0.919 (0.015)	0.695 (0.040)	< 0.001
$r(d; p)$	0.553 (0.049)	0.420 (0.040)	< 0.001
Individual-level means	0.638 (0.036)	0.446 (0.041)	< 0.001

Notes: The standard errors (shown in parentheses) are block-bootstrapped (individual and repeated game level) with 5,000 repetitions, which is used to calculate p-values. The null hypothesis is that the values are identical across the two treatments.

Table 13:
Retaliation Intensities

	Mean	*S.E.*	*p-value*
$r(c;0.9)-r(d;0.9)$	0.374	0.044	0.002+
Individual-level means	0.281	0.030	0.127+
$r(c;0.7)-r(d;0.7)$	0.299	0.031	< 0.001+
Individual-level means	0.249	0.032	< 0.001+
$(r(c;0.9)-r(d;0.9))-(r(c;0.7)-r(d;0.7))$	0.075	0.043	0.156
Individual-level means	0.032	0.043	0.458

Notes: The standard errors are block-bootstrapped (subject and repeated game level) with 5,000 repetitions, which is used to calculate p-values.
+ The hypothesis tests for the comparison to the value implied by the standard theory ($w(p)$), which is 0.235 in the high accuracy treatment and 0.470 in the low accuracy treatment. The null hypothesis is that the mean is identical to the implied value.

Table 14:
Fixed-Effect Model Regression Results on the Experience Effect on Overall Cooperation Rates

	p = 0.9	**p = 0.7**
RG2	0.052 (0.045)	-0.058** (0.044)
RG3	0.075* (0.042)	0.009 (0.044)
RG4	0.102*** (0.031)	0.054 (0.041)
RG5	0.127***(0.034)	0.077*(0.043)
Observations	8,400	8,400
R2	0.016	0.011

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p<0.1$, $**p<0.05$, $***p<0.01$. The coefficient on *RG4* does not differ significantly from that on *RG5* in the both treatment (F-test, $p=0.159$ for the high accuracy treatment, and $p=0.553$ in the low accuracy treatment). The coefficient on *RG3* differs (marginally) significantly from that on *RG5* in both treatments (F-test, $p=0.090$ in the high accuracy treatment, and $p=0.066$ in the low accuracy treatment).

Table 15:
Maximum Likelihood Estimates of Individual Strategies

	p = 0.9	S.E.	p = 0.7	S.E.
ALL-C (g-TFT-1-1)	0.094	(0.086)	0.173***	(0.052)
g-TFT-1-0.75	0.137	(0.105)	0.054	(0.034)
g-TFT-1-0.5	0.124	(0.081)	0.067	(0.047)
g-TFT-1-0.25	0.013	(0.027)	0.003	(0.027)
TFT (g-TFT-1-0)	0	(0)	0.036	(0.030)
Random-0.75 (g-TFT-0.75-0.75)	0.075	(0.052)	0	(0)
g-TFT-0.75-0.5	0.035	(0.032)	0.177**	(0.086)
g-TFT-0.75-0.25	0	(0)	0.083*	(0.047)
g-TFT-0.75-0	0	(0)	0	(0)
Random-0.5 (g-TFT-0.5-0.5)	0.015	(0.017)	0.010	(0.014)
g-TFT-0.5-0.25	0	(0)	0	(0)
g-TFT-0.5-0	0	(0)	0	(0)
Random-0.25 (g-TFT-0.25-0.25)	0	(0)	0.071**	(0.032)
g-TFT-0.25-0	0	(0)	0.	(0.007)
ALL-D (g-TFT-0-0)	0	(0)	0.054*	(0.032)
g-TF2T-0.75	0	(0)	0.072	(0.054)
g-TF2T-0.50	0.273**	(0.134)	0	(0)
g-TF2T-0.25	0	(0)	0	(0)
TF2T (g-TF2T-0)	0	(0)	0	(0)
TF3T	0.055	(0.073)	0	(0)
g-2TFT-0.75	0	(0)	0	(0)
g-2TFT-0.50	0.018	(0)	0	(0)
g-2TFT-0.25	0	(0)	0	(0)
2TFT (g-2TFT-0)	0	(0)	0	(0)
2TF2T	0.033	(0.034)	0.018	(0.033)
Grim	0	(0)	0.019	(0.023)
Grim-2	0	(0)	0.010	(0.013)
Grim-3	0.109	(0.079)	0.146**	(0.057)
Gamma	0.304***	(0.025)	0.332***	(0.063)

Notes: The standard errors (shown in parentheses) are cluster-bootstrapped (individual-level) with 100 repetitions, which is used to calculate p-values. $*p < 0.1$, $**\ p < 0.05$, $***p < 0.01$.

Table 16:
Frequency of g-TFT and Retaliation Intensities (RI)

	p = 0.9	**p = 0.7**
Family of g-TFT	51.2%	72.7%
Excluding signal non-contingent, zero RI strategies (ALL-C, ALL-D, Random-*r*)	32.8%	42.0%
Mean RI conditional on g-TFT	0.251	0.235
Multi-round punishing strategies	16.0%	20.2%
Lenient strategies	47.0%	24.6%

Notes: The equilibrium retaliation intensities are 0.235 in the high accuracy treatment and 0.470 in the low accuracy treatment.

Appendix A: The Proof of the Proposition

Selecting $s_i = C$ instead of D costs player i g in the current round, whereas in the next round she (or he) can gain $1+g$ from the opponent's response with probability $pr(c)+(1-p)r(d)$ instead of $(1-p)r(c)+pr(d)$. Since she must be incentivized to select both actions C and D (belief-free nature), indifference between these action choices must be a necessary and sufficient condition:

$$-g+\delta(1+g)\{pr(c)+(1-p)r(d)\}=\delta(1+g)\{(1-p)r(c)+pr(d)\},$$

or, equivalently,

$$g=\delta(1+g)(2p-1)\{r(c)-r(d)\},$$

implying (1). Since $r(c)-r(d)\leq 1$, the inequality $w(p)\leq 1$ must hold.

Q.E.D.

Appendix B: Derivation of Likelihood

The likelihood function in SFEM frameworks is derived as follows. The choice probability of subject i employing strategy s in round r of repeated game k, given the history of her past choices and signals obtained from the opponent up to the round, is defined as

(B. 1) $$P_{ikr}(s)=\frac{1}{1+\exp(-1/\gamma)},$$

if the observed choice is matched with the predicted choice by strategy s given the history up to the round. Otherwise, the choice is classified as an implementation error, and the choice probability is

(B. 2) $$P_{ikr}(s)=\frac{1}{1+\exp(1/\gamma)},$$

where γ captures the probability of the implementation error.

The likelihood of subject i employing strategy s is

$$P_i(s)=\prod_k \prod_r P_{ikr}(s).$$

In the SFEM framework, the likelihood of subject i over all strategies is a finite mixture of $P_i(s)$ over the entire strategy set. We denote the frequency of occurrence of strategy s by $P(s)$. Then, the log likelihood of the MLE is

$$LH = \sum_i \ln\left(\sum_s P(s)P_i(s)\right).$$

The choice probabilities in (B. 1) and (B. 2) are defined over deterministic strategies. Since the list of strategies considered in our SEFM (Table 7) includes stochastic strategies, the choice probabilities should be extended to cover stochastic cases. Following Fudenberg, Rand, and Dreber (2012), the choice probabilities (B. 1) and (B. 2) are extended to cover stochastic strategies, as follows:

(B. 3)
$$P_{ikr}(s) = s_{ikr}\left(\frac{1}{1+\exp(-1/\gamma)}\right) + (1-s_{ikr})\left(\frac{1}{1+\exp(1/\gamma)}\right)$$

if the observed choice is C,

(B. 4)
$$P_{ikr}(s) = (1-s_{ikr})\left(\frac{1}{1+\exp(-1/\gamma)}\right) + s_{ikr}\left(\frac{1}{1+\exp(1/\gamma)}\right)$$

if the observed choice is D,

where s_{ikr} is the probability of playing C in stochastic strategy s given the history up to the round. Observe that the new formulations of the choice probabilities (B. 3) and (B. 4) are reduced to the previous definition (B. 1) and (B. 2) when s_{ikr} takes a value of either 1 or 0 as deterministic choices.

The standard error of the MLE is computed through a cluster-bootstrap (individual-level) with 100 resamples, which is also used to perform the hypothesis tests presented in Section 8.

Online Appendix C-I

Accuracy and Retaliation in Repeated Games with Imperfect Private Monitoring: Experiments

Appendix C: Order Effect

In this section, we document the order effect of the treatment observed in our experiment. However, note that the order effect should not appear prominently at the aggregate level since the treatments in our experiments are counterbalanced with reverse ordered treatment.

[TABLE C. 1 HERE]

Table C.1 displays the regression results of the order effect on overall cooperation rates. The regressor *Low First* is a dummy variable that takes a value of 1 if the subject's first treatment is the low accuracy one. In the high accuracy treatment, the coefficient on *Low First* is marginally significantly positive ($p < 0.1$). This indicates that subjects who receive the low accuracy treatment first could be more cooperative than other subjects. Similarly, the coefficient in the low accuracy treatment is positively significant ($p < 0.05$), indicating that subjects tend to be more cooperative when the first treatment is the low accuracy treatment.

[TABLE C. 2 HERE]

[TABLE C. 3 HERE]

Table C.2 and Table C.3 display the regression results on the order effect on the signal-contingent cooperation rates. Similarly, the coefficients on *Low First* tend to be positive. However, a significantly large one exists only in the case that is contingent on the good signal in the low accuracy treatment.

[TABLE C. 4 HERE]

Although we find some order effects on the overall cooperation rates and signal contingent cooperation rates, we do not find any significant order effect on the retaliation intensities. Table C.4 displays the regression results on the order effect on the retaliation intensities. None of the coefficients on the joint effect on *Signal* and *Low First* are significantly different from zero in both treatments. The level of cooperation could be affected by the treatment order. However, the strengths of retaliations are not.

Table C.1:
Regression Results on the Order Effect on Overall Cooperation Rates

	p = 0.9	p = 0.6
Const.	0.631*** (0.035)	0.303*** (0.032)
Low First	0.088* (0.050)	0.107** (0.050)
Observations	8,864	9,144
R2	0.009	0.0013

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p<0.1$, $**p<0.05$, $***p<0.01$.

Table C.2:
Regression Results on the Order Effect on Cooperation Rates Contingent on Good Signals

	p = 0.9	p = 0.6
Const.	0.827*** (0.024)	-0.376*** (0.041)
Low First	0.049 (0.034)	0.122** (0.061)
Observations	5,453	4,265
R2	0.005	0.015

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p<0.1$, $**p<0.05$, $***p<0.01$.

Table C.3:
Regression Results on the Order Effect on Cooperation Rates Contingent on Bad Signals

	p = 0.9	p = 0.6
Const.	0.327*** (0.036)	0.236*** (0.029)
Low First	0.040 (0.054)	0.078 (0.049)
Observations	3,087	4,555
R2	0.017	0.077

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table C.4:
Regression Results on the Order Effect on Retaliation Intensities

	p = 0.9	p = 0.6
Const.	0.322*** (0.036)	0.239*** (0.030)
Signal	0.506 ***(0.034)	0.144 *** (0.032)
Low First	0.050 (0.054)	0.068 (0.049)
Signal: Low First	-0.005 (0.47)	0.035 (0.046)
Observations	8,540	8,820
R2	0.271	0.038

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Appendix D: Screens in Constant Random Termination

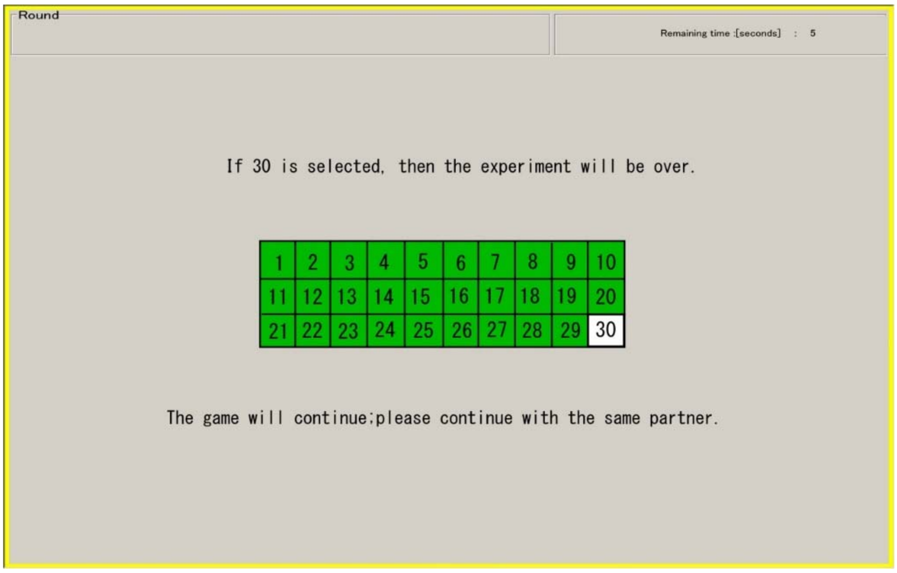


Figure D.1: Screen when the repeated game continues

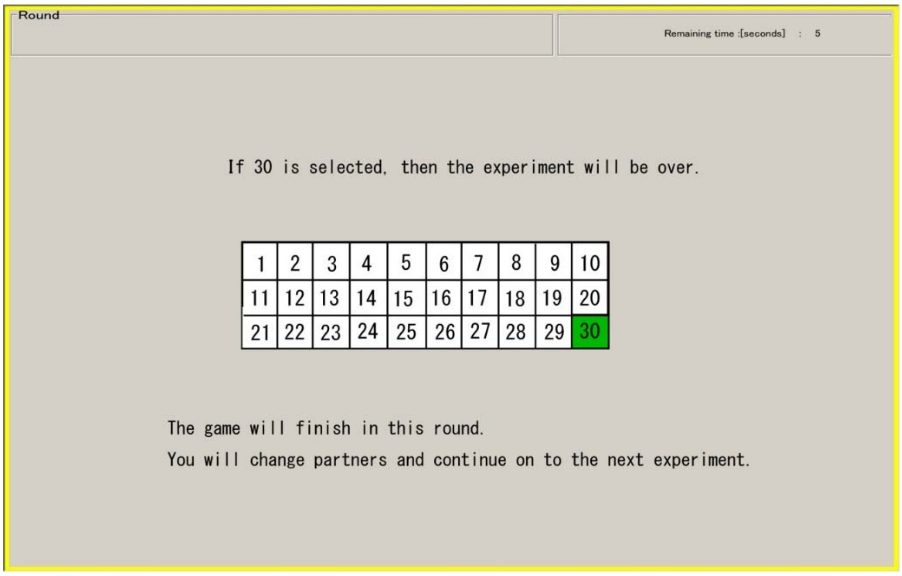


Figure D.2: Screen when the repeated game is terminated

Appendix E: Impact of Experience

Several existing studies have reported that the frequency of cooperation changes as people experience playing repeated games (see discussion in Dal Bó and Fréchette, 2016). The overall cooperation rates might rise with the experience of repeated games when the experimental parameters are conducive to cooperation. Although the sizes are not large, non-negligible learning effects are observed in the study. We document the learning effects in detail in this section.

To examine the impact of the experience of repeated games on overall cooperation rates, we perform the following reduced-form, linear regression analysis; we regress the action choices on the two explanatory variables, *RG2* and *RG3*. The dummy variable *RG2* (*RG3*) takes a value of 1 if the choice is made in the second (third) repeated game in each treatment. Moreover, to further examine the change of the cooperation rate within a repeated game, we include the dummy variable *First 14 Rd*, which takes a value of 1 if the choice is made in the first 14 rounds of each repeated game. The regression model is a fixed-effect model in which the individual heterogeneity in the tendency to adopt cooperative choices is controlled by individual fixed effects.

[TABLE E.1 HERE]

Table E.1 displays the regression results. In the high accuracy treatment, the coefficient on the second repeated game is 0.029 (insignificant, $p = 0.355$), and the coefficient on the third repeated game is 0.105 (significant, $p < 0.001$). These positive values in the regression coefficients indicate that our subjects tend to cooperate more as they gain experience, indicating that the welfare of the two players is improved by experience. However, the size is at most 11%, which is not remarkably large.

In the low accuracy treatment, the coefficient on the second repeated game is −0.071 (significant, $p < 0.05$), and the coefficient on the third repeated game is −0.118 (significant, $p < 0.001$). Contrary to the case in the high accuracy treatment, our subjects tend to become less cooperative as they gain experience, even when the experimental parameters are conducive to cooperation, indicating that welfare worsens with experience. However, the sizes of the effects of experience are at most 12%, which is again not remarkably large.

As for the effect within a repeated game, the coefficients on *First 14 Rd* are statistically significant and larger than 0 in both treatments ($p < 0.001$ for both treatments), although the sizes are at most 8%. Our subjects tend to become less cooperative as the rounds proceed in each repeated game, but the effect is small.

These results indicate that, although there are some experience effects on action choices, the sizes of the effects are not remarkably large in our data. Given that the three repeated games include approximately 90 rounds, the subjects' learning is slow. Additionally, we perform an identical analysis on the signal contingent cooperation rates and find qualitatively similar results (Tables E.2 and E.3).

Studies on repeated games frequently report that cooperation often increases with experience when the experimental parameters are favorable for cooperation. Although the learning effects are not large in our experiment, similar behavior is observed in the study. Our subjects become more cooperative adaptively as they gain experience in the high accuracy treatment. Perhaps this is due to the observed retaliations that are stronger than implied by the standard theory. The better response under stronger retaliations is to be more cooperative to avoid stronger punishments from defecting. On the contrary, in the low accuracy treatment, our subjects become less cooperative with experiences adaptively. This might be due to our subjects' retaliation intensity, which is weaker than implied by the standard equilibrium. In such a situation, a better response is to be more defective to exploit the weaker punishments.

Although examining the causal relationship between learning and retaliation intensity is an interesting issue, however, we cannot pursue the causal relationship in the study, since our experimental design does not aim to identify it. However, the learning effects observed in the study are consistent with the explanations of the standard theory.

[TABLE E.2 HERE]

[TABLE E.3 HERE]

Since the primary focus of the study is retaliation intensity, we also investigate the effect of experience on retaliation intensities. Here, we perform a similar, reduced-form regression analysis, regressing the action choices on the dummy variable *Signal*, which takes a value of 1 if the signal is good. The coefficient on the dummy variable captures the contrast between the cooperation rate contingent on the good signal and the cooperation rate contingent on the bad signal, which is the retaliation intensity. To examine the experience effects on retaliation intensities across repeated games and within a repeated game, we add the cross-product terms with *RG2*, *RG3*, and *First 14 Rd* in the set of explanatory variables. Again, the regression model is a fixed-effect model in which individual heterogeneity in the tendencies to adopt cooperative choices is controlled by individual fixed effects.

[TABLE E.4 HERE]

Table E.4 displays the regression results. None of the coefficients on the joint effect of *Signal* and repeated games (*RG2* and *RG3*) is significantly different from 0 in both treatments, implying that the retaliation intensities do not differ either in the second repeated games or in the third repeated games from those in the first repeated games. Furthermore, the coefficient on the cross-product of *Signal* and *RG2* does not differ from that on the cross-product of *Signal* and *RG3* in both treatments ($p = 0.702$ for the high accuracy treatment and $p = 0.672$ for the low accuracy treatment). These results jointly indicate that the retaliation intensities are stable across repeated games.

As for the within repeated game difference, only the coefficient on the joint effect of *Signal* and *First 14 Rd* in the high accuracy treatment significantly differs from 0, although the size of the effect is approximately 5%. Our subjects tend to rely on stronger retaliation intensities as the rounds proceed in a repeated game, but the difference is not remarkably large.

The results here indicate that retaliation intensities do not change significantly as our subjects gain experience, contrary to the case of overall cooperation rates and signal contingent cooperation rates.

Table E.1:
Fixed-Effect Model Regression Results on the Experience Effect on Overall Cooperation Rates

	p = 0.9	p = 0.6
RG2	0.029 (0.031)	-0.071** (0.027)
RG3	0.105*** (0.031)	-0.118*** (0.029)
First 14 Rd	0.062*** (0.014)	0.080*** (0.015)
Observations	8,864	9,144
R2	0.021	0.025

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. The coefficient on *RG3* is significantly larger than that on *RG2* in the high accuracy treatment (F-test, $p = 0.016$). The coefficient on *RG3* is significantly smaller than that on *RG2* in the low accuracy treatment (F-test, $p = 0.027$).

Table E.2:
Fixed-Effect Model Regression Results on the Experience Effect on Cooperation Rates Contingent on Good Signals

	p = 0.9	p = 0.6
RG2	0.014 (0.016)	-0.055* (0.029)
RG3	0.076*** (0.021)	-0.104*** (0.034)
First 14 Rd	0.018 (0.011)	0.073*** (0.018)
Observations	5,453	4,265
R2	0.005	0.021

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table E.3:
Fixed-Effect Model Regression Results on the Experience Effect on Cooperation Rates Contingent on Bad Signals

	p = 0.9	p = 0.6
RG2	0.025 (0.030)	-0.070** (0.030)
RG3	0.052 (0.036)	-0.123*** (0.029)
First 14 Rd	0.061*** (0.018)	0.070*** (0.016)
Observations	3,087	4,555
R2	0.009	0.027

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

Table E.4:
Fixed-Effect Model Regression Results on the Experience Effect on Retaliation Intensities

	p = 0.9	**p = 0.6**
Signal	0.392*** (0.030)	0.118*** (0.025)
Signal: RG2	-0.015 (0.027)	0.020 (0.022)
Signal: RG3	0.007 (0.038)	0.034 (0.027)
Signal: First 14 Rd	-0.048** (0.019)	0.005 (0.018)
RG2	0.029 (0.034)	-0.077** (0.030)
RG3	0.070* (0.039)	-0.130*** (0.029)
First 14 Rd	0.068*** (0.018)	0.072*** (0.016)
Observations	8,540	8,820
R2	0.021	0.025

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. The coefficient on the cross-product term *Signal*: *RG2* is not significantly different from that on the cross-product term *Signal*: *RG3* in the high accuracy treatment (F-test, $p = 0.702$). The coefficient on the cross-product term *Signal*: *RG2* is not significantly different from that on the cross-product term *Signal*: *RG3* in the low accuracy treatment (F-test, $p = 0.672$).

Appendix F: The Validity of SFEM with a Large Set of Generous-TFT Strategies

Our case is the first study in which a large set of mixed strategies are included in the strategy list. Simulation exercises by Fudenberg, Rand, and Dreber (2012) include mainly deterministic strategies with a few g-TFT strategies in their strategy list to investigate whether the SFEM could distinguish memory-1 strategies from longer memory strategies. Similarly, we perform simulation exercises to investigate the efficacy of our SFEM, whose strategy list includes an extensive set of g-TFT strategies in our experimental environment.

Related to the main claims of the study, the aims of the simulation exercises here are as follows:

1. Observing that the SFEM correctly distinguishes long-memory (lenient) strategies from strategies in a class of g-TFT strategies (i.e., memory-1 strategy) and is able to estimate the fraction of g-TFT strategies correctly.
2. Observing that the SFEM correctly estimates the mean of retaliation intensities among g-TFT players.

In the following simulations, mimicking our experimental environments, we generate 108 subjects, and each subject experiences three interactions. Each interaction includes 28 rounds. Partners randomly alter across interactions. The SFEM strategy list employed in the following simulations is the same as that employed in the main study (Table 7). We repeated the data generation and estimation 100 times for each simulation, and we report means and standard deviations of the estimates.

A. High accuracy case

Here we set up the following three types of players:

a. G-TFT-1-0.5, which has the largest individual share in our SFEM estimates in the high accuracy treatment.
b. TFT, to examine whether the SFEM correctly distinguishes g-TFT-1-0.5 from TFT.

c. TF2T, to examine whether the SFEM correctly distinguishes the long-memory (lenient) strategy from strategies in a class of g-TFT strategies (i.e., memory-1 strategy).

In this simulation, one-third of the total subjects (i.e., 36 subjects) play g-TFT-1-0.5, another one-third play TFT, and the rest play TF2T. The signal accuracy is set to the higher accuracy (i.e., $p = 0.9$).

[TABLE F.1 HERE]

Table F.1 summarizes the results of the SFEM. As shown, each estimate for g-TFT-1-0.5, TFT, and TF2T is approximately one-third, indicating that the SFEM successfully distinguishes the mixed strategy from the deterministic strategies and distinguishes the long-memory (lenient) strategy from the strategies in a class of g-TFT strategies.

Based on the accurate individual estimates above, the total fraction of the strategies in a class of g-TFT strategies is correctly estimated, which is 0.6676 (true: two-thirds consisting of g-TFT-1-0.5 and TFT). Additionally, the mean of retaliation intensities among g-TFT strategies is correctly estimated. The estimate is 0.7518 (true: 0.7500).

B. Low accuracy case

Here we set up the following three types of players:

a. ALL-D, which has the largest individual share in our SFEM estimates in the low accuracy treatment.
b. G-TFT-0.625-0.375, which has the largest and significant individual share among g-TFT strategies other than ALL-D.
c. TF2T, to examine whether the SFEM successfully distinguishes the long-memory (lenient) strategy from strategies in a class of g-TFT strategies (i.e., memory-1 strategy).

In this simulation, one-third of the total subjects play ALL-D, another one-third play g-TFT-0.625-0.375, and the rest play TF2T. The signal accuracy is set to the lower accuracy (i.e., $p = 0.6$).

[TABLE F.2 HERE]

Table F.2 summarizes the results of the SFEM. Each estimate for ALL-D and TF2T is approximately one-third. The SFEM accurately distinguishes the deterministic strategies from the mixed strategy and distinguishes the long-memory (lenient) strategy from the strategies in a class of g-TFT strategies.

The estimate for g-TFT-0.625-0.375 has some bias. The estimate is 0.2841, which is smaller than the true value by approximately 0.05 (true: one-third). Instead, similar "nearby" g-TFT strategies (g-TFT-0.75-0.375, g-TFT-0.75-0.25, and g-TFT-0.625-0.5) have some small positive estimates (true: 0 for each). However, the largest value is only 0.037 in g-TFT-0.75-0.35.

Since the bias occurs only among "nearby" g-TFT strategies, its effect on the estimate of the total fraction of g-TFT strategies and on the estimate of the mean of retaliation intensities should be at most marginal. Indeed, the total share of the strategies in a class of g-TFT strategies is accurately estimated, which is 0.6667 (true: two-thirds consisting of g-TFT-0.625-0.375 and ALL-D). Additionally, the mean of retaliation intensities among g-TFT players is approximately correctly estimated. The estimate is 0.1335, which is slightly larger than the true value (0.1250), by approximately 0.008.

Table F.1:
Mean of SFEM Estimates for Simulated Data in High Accuracy Case (p = 0.9)

	True	Estimate
Generous-TFT-1-0.625	0	0.001 (0.003)
Generous-TFT-1-0.5	1/3	0.323 (0.012)
Generous-TFT-1-0.37.5	0	0.007 (0.011)
Generous-TFT-1-0.25	0	0.003 (0.004)
TFT	1/3	0.333 (0.000)
TF2T	1/3	0.333 (0.000)

Other strategy (each)	0	Less than 0.001

Notes: Standard deviations in parenthesis.

Table F.2:
Mean of SFEM Estimates for Simulated Data in Low Accuracy Case (p = 0.6)

	True	**Estimate**
Generous-TFT-0.75-0.5	0	0.001 (0.006)
Generous-TFT-0.75-0.375	0	0.037 (0.017)
Generous-TFT-0.75-0.25	0	0.006 (0.009)
Generous-TFT-0.625-0.5	0	0.004 (0.007)
Generous-TFT-0.625-0.375	1/3	0.284 (0.023)
Generous-TFT-0.625-0.25	0	0.001 (0.005)
ALL-D	1/3	0.333 (0.000)
TF2T	1/3	0.333 (0.000)
Other strategy (each)	0	Less than 0.001

Notes: Standard deviations in parenthesis.

Appendix G: Robustness Checks of Our Strategy Estimation

In this part of the appendix, we discuss the robustness of the SFEM estimates. In the main text, we use all three repeated games of each treatment in our estimation. Here, we demonstrate that the estimation results show almost no changes, even using only the final two repeated games in each treatment (Tables G.1 and G.2).

The mean retaliation intensity in the high accuracy treatment decreases slightly in the final two repeated games, which is closer to the value implied by the g-TFT equilibria (0.235), and which might suggest that our subjects learn optimal retaliation intensities by experience in the high accuracy treatment.

Table G.1:
Aggregated Estimates for High Accuracy Case (p = 0.9)

	All	Final 2
Cooperative strategies	83.9%	88.3%
Family of g-TFT (including zero RI strategies)	76.8%	75.2%
Equilibrium RI	5.6%	7.6%
Stronger RI	54.4%	49.0%
Weaker RI (including zero RI strategies)	17.7%	20.8%
Mean RI conditional on g-TFT (including zero RI strategies)	0.426	0.360
Lenient strategies	21.4%	21.6%

Table G.2:
Aggregated Estimates for Low Accuracy Case (p = 0.6)

	All	Final 2
Cooperative strategies	34.7%	37.6%
Family of g-TFT (including zero RI strategies)	96.2%	96.5%
Equilibrium RI	2.7%	2.4%
Stronger RI	0.0%	0.0%
Weaker RI (including zero RI strategies)	93.5%	94.0%
Mean RI conditional on g-TFT (including zero RI strategies)	0.148	0.135
Lenient strategies	3.7%	2.5%

Appendix H:
Experimental Instruction and Computer Screen Images
(October 2006, Translation from Japanese into English)

1. Experimental Instruction

Please make sure all the contents are in your envelope. The envelope should have the following items.

1. Pen – 1
2. Instruction – 1 copy
3. Printed computer screen images – 1 copy
4. Bank transfer form – 1 sheet
5. Scratch paper – 1 sheet

If you are missing any item, please raise your hand quietly. We will collect the items at the end of all the experiments, except for the scratch paper, which you can keep.

Please look at the instructions (this material). You will be asked to make decisions at a computer terminal. You will earn "points" according to your performance in the experiments. The points will be converted into monetary rewards at the exchange rate of 0.6 yen per point, which will be paid in addition to the participation fee (1,500 yen). The total amount of money you will receive from the experiments is

the number of points earned × 0.6 yen + participation fee of 1,500 yen.

Any communication with other participants (i.e., conversation or exchange of signals) is not allowed during the experiments; if you violate this rule, you may be asked to leave the experiments. Furthermore, you are not allowed to leave in the middle of the experiments unless an experimenter allows or asks you to do so. Please turn off your cell phones during the experiments.

Outline of Experiments

We will conduct six experiments across two sessions. Each session includes three experiments with one practice experiment preceding them. The six experiments are independent of each other; the records of one experiment are not transferred to the other experiments. The experiments are conducted via a computer network. You are asked to make decisions at a computer terminal and interact with other participants through the computer network.

All the participants will be divided into pairs in each experiment. The pairs are selected randomly by the computer.

Each experiment consists of several rounds (i.e., Rounds 1, 2, 3, etc.). Later, we will explain the rule that decides the number of rounds conducted in each experiment. In each round, you are asked to choose one of two alternatives, which will also be explained below.

Please raise your hand quietly if you have any questions.

Decision Making

You will be asked to choose either A or B in each round. Your partner will also be asked to choose either A or B. Please look at the table.

You \ Your partner	A	B
A	60 60	5 70
B	70 5	15 15

The table summarizes the points you and your partner earn according to the combination of choices made by the two players. The characters in the left column marked in red indicate your choice, which is either A or B. The characters in the top row marked in light blue indicate the choice of your partner, which is also either A or B. In each cell, the numbers in red on the left side indicate the points you earn, and the numbers in light blue on the right side indicate the points your partner earns.

If both you and your partner select A,

both you and your partner earn 60 points.

If you select A and your partner selects B,

you earn 5 points, and your partner earns 70 points.

If you select B and your partner selects A,

you earn 70 points, and your partner earns 5 points.

If both you and your partner select B,

both you and your partner earn 15 points.

Please look at the table carefully and ensure that you understand how the points will be awarded to you and your partner according to the choices made by the two players. Your earnings depend not only on your choice but also on the choice of your partner. Similarly, your partner's earnings depend on your choice as well as her own.

Please raise your hand quietly if you have any questions.

Session 1

Session 1 consists of three experiments, numbered 1, 2, and 3. The three experiments follow identical rules and will be conducted consecutively.

Observable Information

You are not allowed to observe whether your partner selected A or B directly. However, you will receive signal a or signal b, which has information about your partner's choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to you.

If your partner selects A, you will receive

signal a with a 90% chance and signal b with a 10% chance.

If your partner selects B, you will receive

signal b with a 90% chance and signal a with a 10% chance.

In the same way, your partner will not know whether you have selected A or B. However, your partner will receive signal a or signal b, which has information about your choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to your partner.

If you select A, your partner will receive

signal a with a 90% chance and signal b with a 10% chance.

If you select B, your partner will receive

signal b with a 90% chance and signal a with a 10% chance.

The signal you receive and the signal your partner receives are decided independently and have no correlation. Furthermore, the computer determines the signals independently in each round.

We refer to the stochastic rules for this signal-generating process as **signaling with 90% accuracy.**

Please raise your hand quietly if you have any questions.

Number of Rounds

The number of rounds in each experiment will be determined randomly. At the end of each round, the computer will randomly select a number from 1 to 30 without replacement, so there is a 1/30 chance of any number being selected by the computer. The number selected by the computer is applied uniformly to all participants.

The experiment will be terminated when the number 30 is selected by chance.

The experiment will continue if any number other than 30 is selected. However, you will notice only that a number other than 30 is selected, instead of seeing the specific number selected by the computer. Then, you will move on to the next round and will be asked to make a decision faced with the same partner.

The probability that the experiment is terminated in each round remains the same, which is 1/30, regardless of the number of rounds (1, 2, 3, etc.). However, the maximum possible number of rounds in an experiment, which is experimentally controlled, is 98.

When Experiment 1 is terminated, you proceed to Experiment 2, and you will be randomly paired with a new partner. When Experiment 2 is terminated, you proceed to Experiment 3, and again, you will be randomly paired with a new partner. Session 1 will be over when Experiment 3 is terminated.

Please raise your hand quietly if you have any questions.

Description of Screens and Operations for Computers

Please look at the booklet with printed computer screen images.

Please look at Screen 1 and Screen 2. Screen 1 displays the screen that will be presented to you during the decision phases. Screen 2 is the screen that will be presented to your partner during the decision phases. Please look at the top left portion of each screen, which indicates that the current round is Round 4. The left portion of Screen 1 displays the information available to you up to the round. The left portion of Screen 2 presented to your partner displays the information available to her up to the round.

You are asked to click with the mouse to select either “A” or “B” in the bottom right portion of the screen. Then, the selection will be confirmed by clicking the “OK” button right below the alternatives.

Next, please look at Screen 3 and Screen 4. Screen 3 presents the results to you. Screen 4 presents the results to your partner. The screens display the situation in which, in Round 4, both you and your partner chose A. Screen 3 shows you that, in Round 4, “your partner’s signal (accuracy: 90%) is b,” indicating to you that the signal you observe about the partner’s choice is “b.” On the other hand, Screen 4 shows your partner that, in Round 4, “your partner’s signal (accuracy: 90%) is a,” indicating to your partner that the signal your partner observes about your choice is “a.” Recall that your partner will observe signal a with a probability of 90% and will observe signal b with a probability of 10% when you choose “A.”

Then, we move on to the lottery screens. Please turn the page and look at Screen 5 and Screen 6, which display the lottery. Any number from 1 to 30 will be randomly selected with an identical probability of occurrence, which is 1/30. Then, a part of the cells turns green according to the number selected. If the number 30 is selected, the cell numbered 30 turns green and the message below explains that the current experiment is terminated.

Otherwise, Screen 5 is shown, in which all the cells numbered 1 to 29 turn green at once (you do not know which number is selected specifically), and the message below explains that the experiment continues with the same partner. Screen 6 is presented when the number 30 is selected, and the cell numbered 30 turns green, indicating that the current experiment is terminated in that round. Again, please make sure that the experiment is terminated when the cell numbered 30 turns green.

Finally, please look at Screen 7. This screen is presented at the end of each experiment. The screen displays the total number of points you earned in the experiment, the average number of points per round, the total number of points your partner earned, and the average number of points per round of your partner. Then, you will be matched with a new partner and move on to the next experiment.

Please raise your hand quietly if you have any questions.

Session 2

Please look at page 6 of your instruction. In Session 2, you will participate in three experiments (namely, Experiments 4, 5, and 6), with a practice experiment preceding them.

The three experiments follow identical rules and will be conducted consecutively. Session 2 proceeds similarly to Session 1, but the signal accuracy of Session 2 is different from that of Session 1. Except for the signal accuracy, the two sessions are identical.

Observable Information

You are not allowed to observe directly whether your partner selected A or B. However, you will receive signal a or signal b, which has information about your partner's choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to you.

If your partner selects A, you will receive

signal a with a 60% chance and signal b with a 40% chance.

If your partner selects B, you will receive

signal b with a 60% chance and signal a with a 40% chance.

In the same way, your partner will not know whether you selected A or B. However, your partner will also receive signal a or signal b, which has information about your choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to your partner.

If you select A, your partner will receive

signal a with a 60% chance and signal b with a 40% chance.

If you select B, your partner will receive

signal b with a 60% chance and signal a with a 40% chance.

The signals you and your partner receive are decided independently and have no correlation. Furthermore, the computer determines the signals independently in each round.

We refer to the stochastic rules for the signal-generating process as **signaling with 60% accuracy.**

Please raise your hand quietly if you have any questions.

Description of Screens and Operations for Computers

Please look at the booklet with printed computer screen images.

Please look at Screen 8 and Screen 9. Screen 8 displays the screen that will be presented to you during the decision phases. The left portion of Screen 8 displays the information available to you up to the round. Please check the current signal accuracy with the message "Signal accuracy for your partner's selection: 60%." Screen 9 is the screen that will be presented to your partner during the decision phases. The left portion of Screen 9 presented to your partner displays the information available to her up to the round.

Please look at Screen 10 and Screen 11 on page 6. Screen 10 presents the results to you.

Screen 11 presents the results to your partner. The screens capture the situation in which, in Round 4, both you and your partner chose A, but you observed signal b and your partner observed signal a. Please make sure that the bottom right portions of Screen 10 and Screen 11 display the signals you and your partner observed, respectively.

The choices you made and the signals you observe about your partner's choices are only available to you and are not available to your partner. Please make sure of this point in Screen 10 and Screen 11.

The screens for the lottery on the continuation of the experiment are identical to those in Session 1, which show numbers 1 to 30. Please refer to Screen 5 and Screen 6. The results screen at the end of the experiment is also identical to that in Session 1 (Screen 7).

Please raise your hand quietly if you have any questions.

Now, all the processes of the experiments have been completed, and all the points awarded to everyone recorded on the computer.

Please answer the questionnaire that will be distributed now.

Take the bank transfer form out of the envelope and fill it out accurately; otherwise, we will not be able to process the payment correctly for you.

Please raise your hand quietly if you have any questions.

Please make sure that you fill out the questionnaire and the bank transfer form correctly.

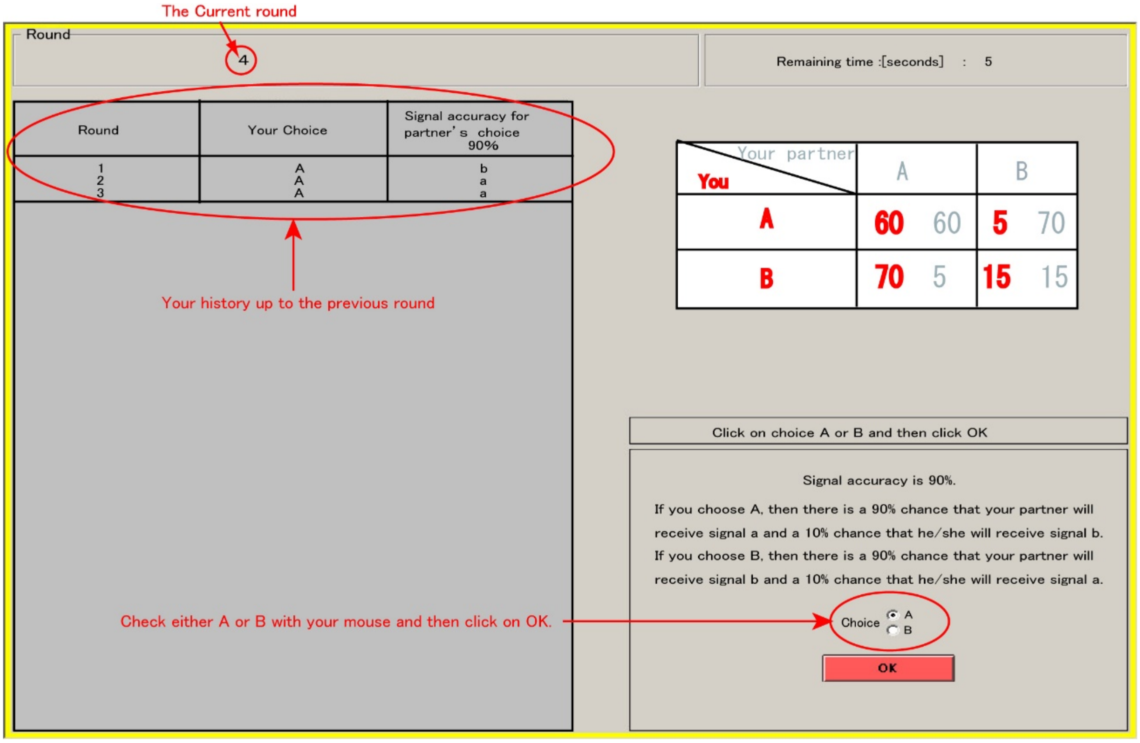
Please raise your hand quietly if you have any questions.

Please put all the documents in the envelope. Please leave the pen and ink pad on the desk. Make sure you take all your belongings with you when you leave.

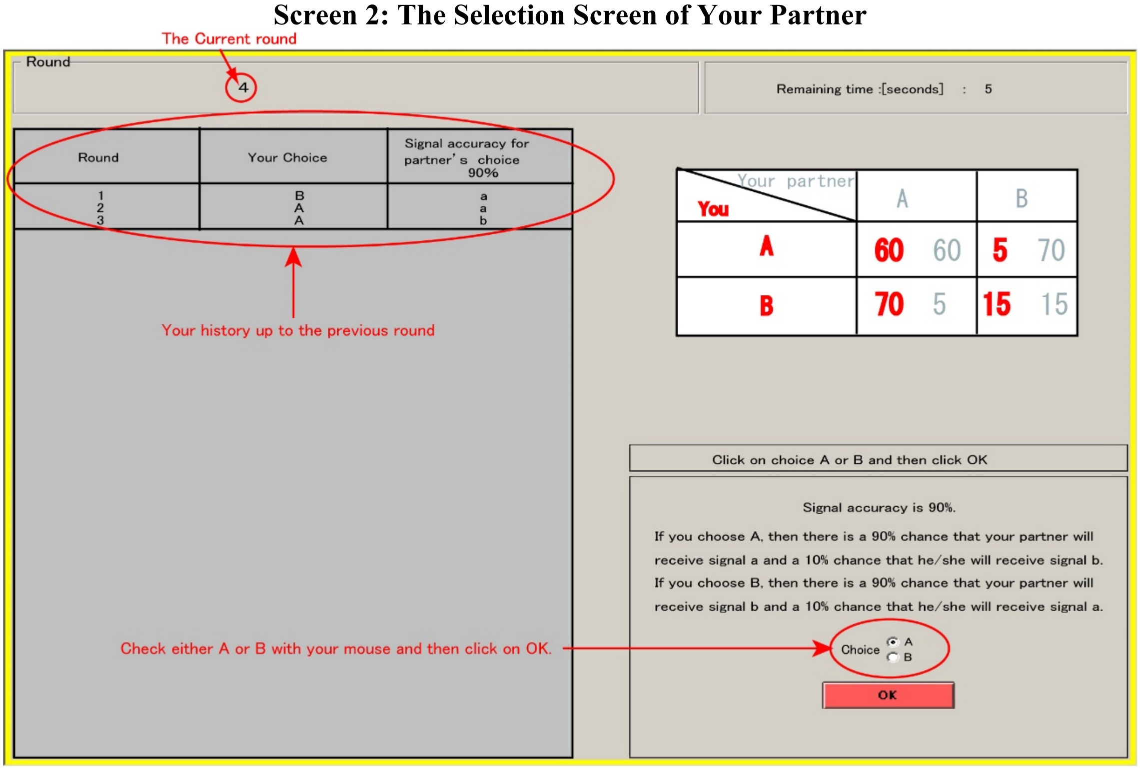
Please do not disclose any details regarding the experiments to anyone until Saturday. Thank you very much for your participation. Please follow the instructions of the experimenters to leave the room.

2. Computer Screen Images

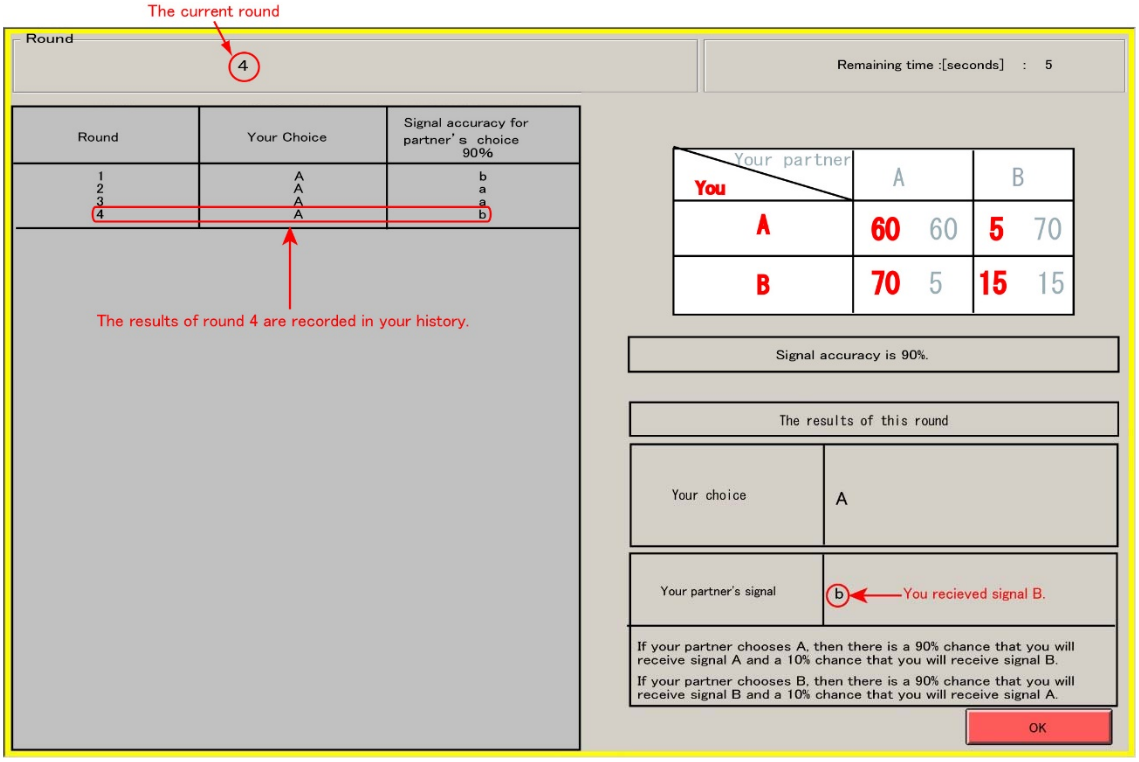
Screen 1: Your Selection Screen



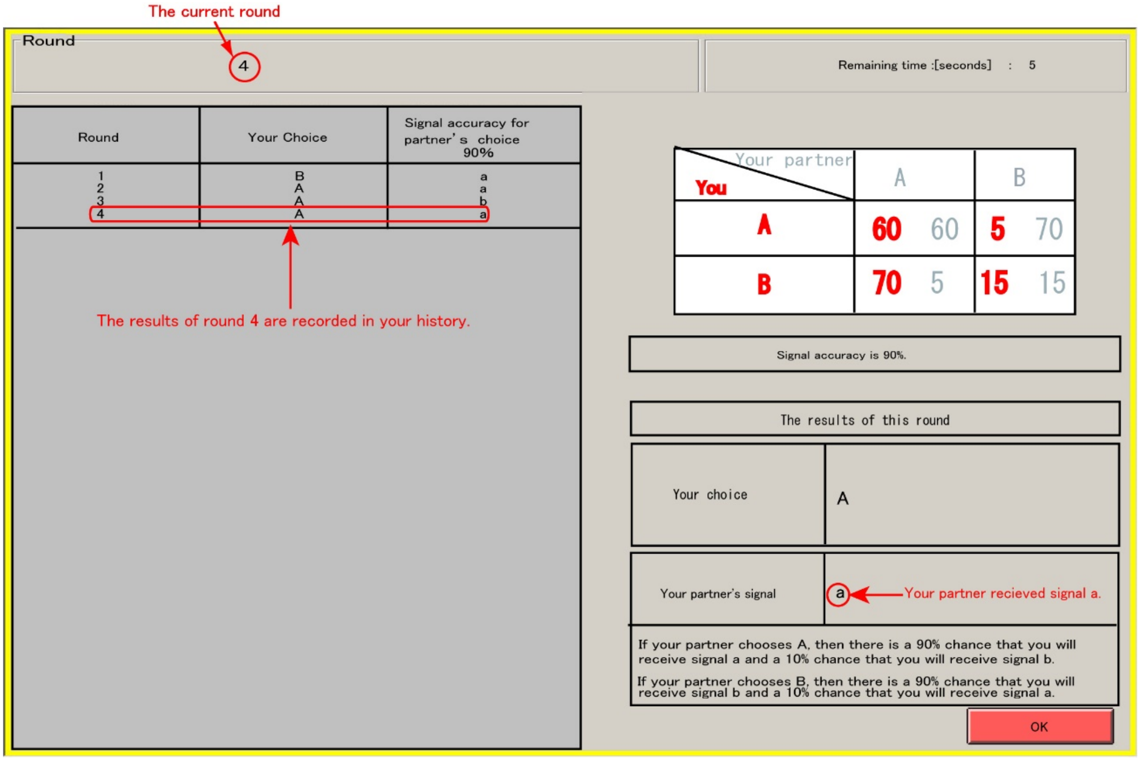
Screen 2: The Selection Screen of Your Partner



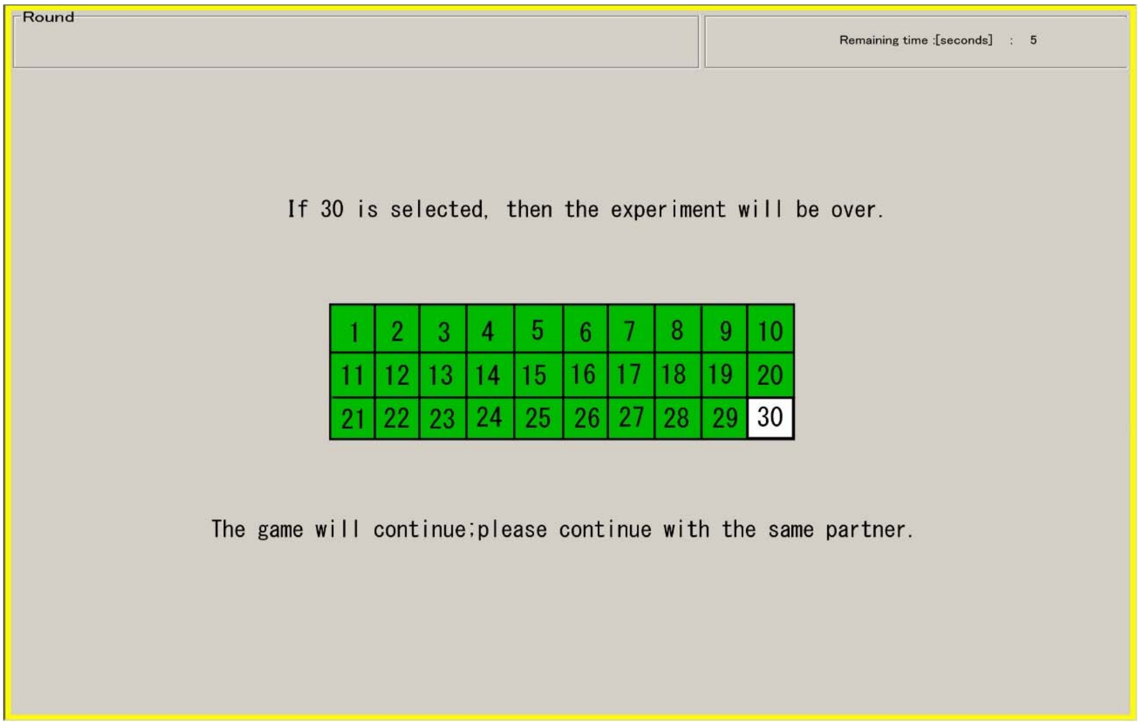
Screen 3: Your Results Screen



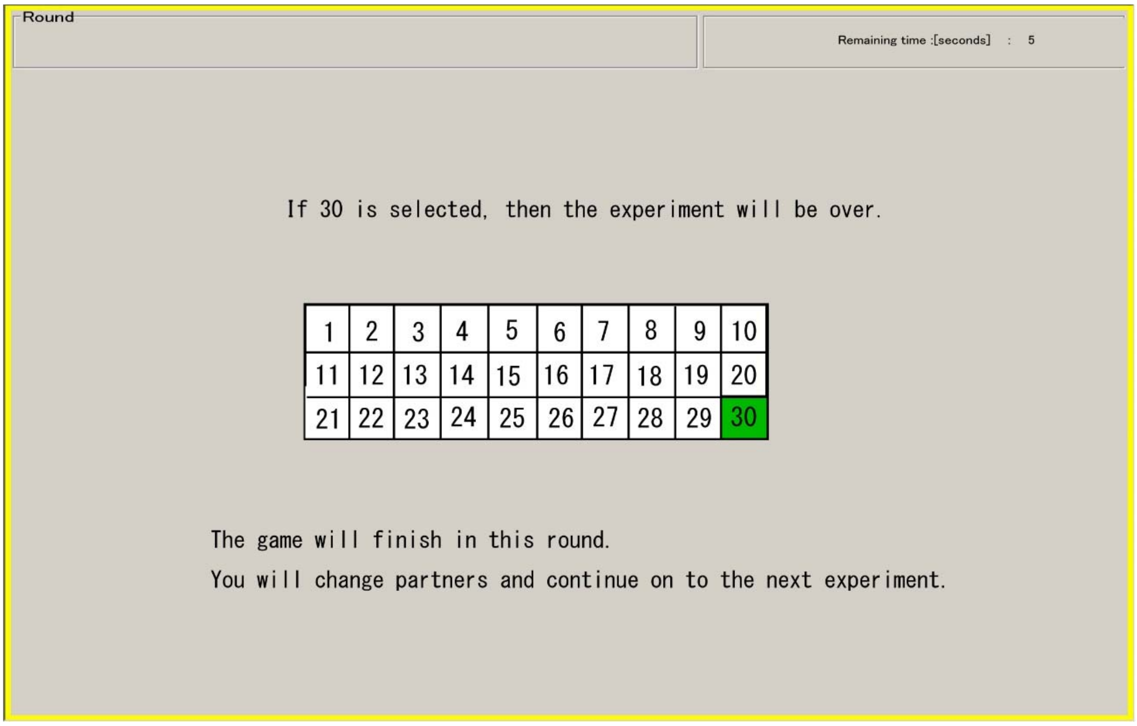
Screen 4: The Results Screen of Your Partner



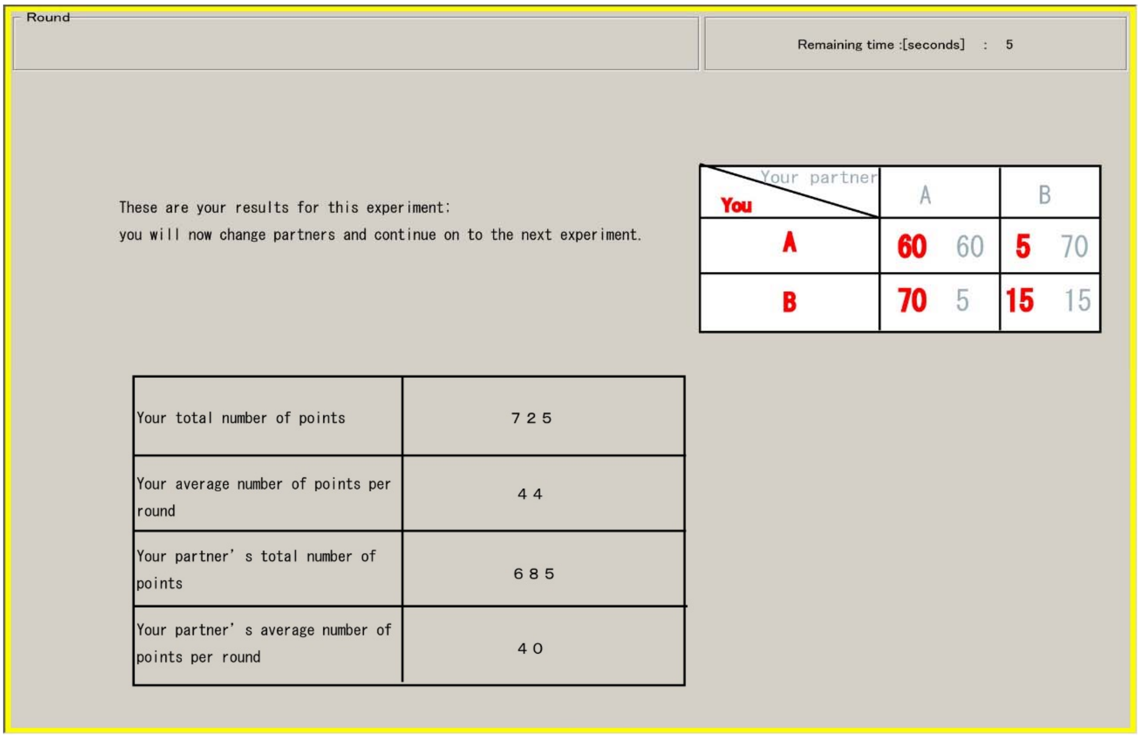
Screen 5: Lottery (experiment continues)



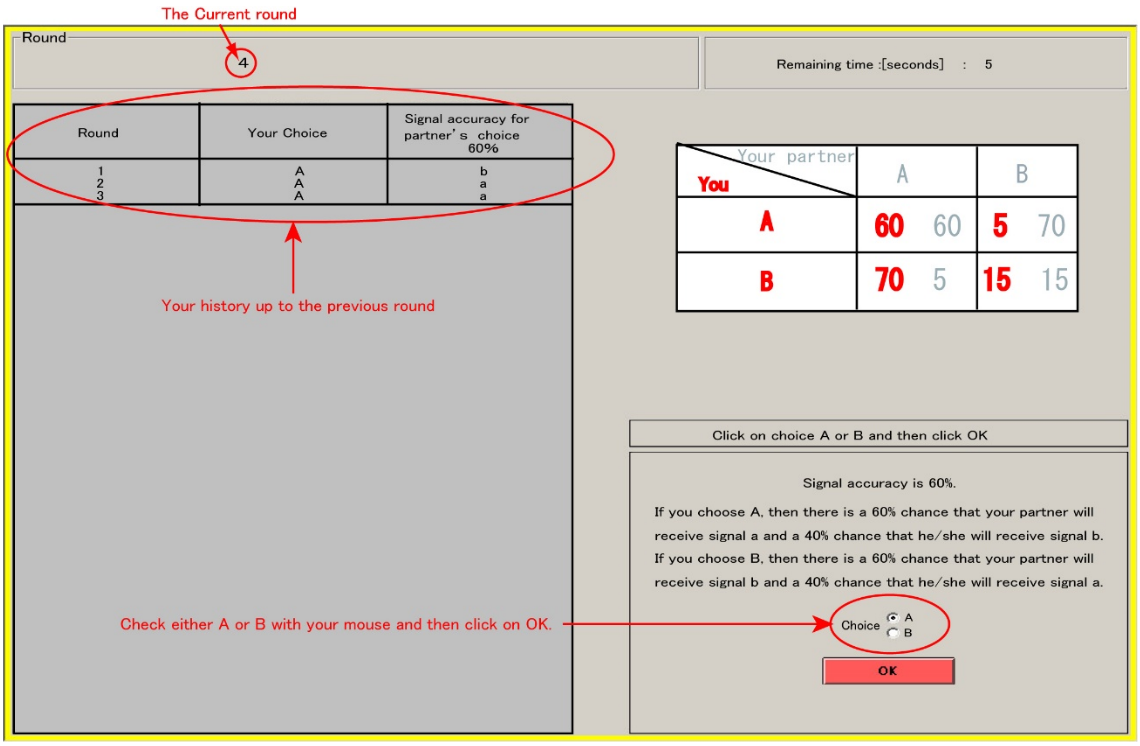
Screen 6: Lottery (experiment is over)



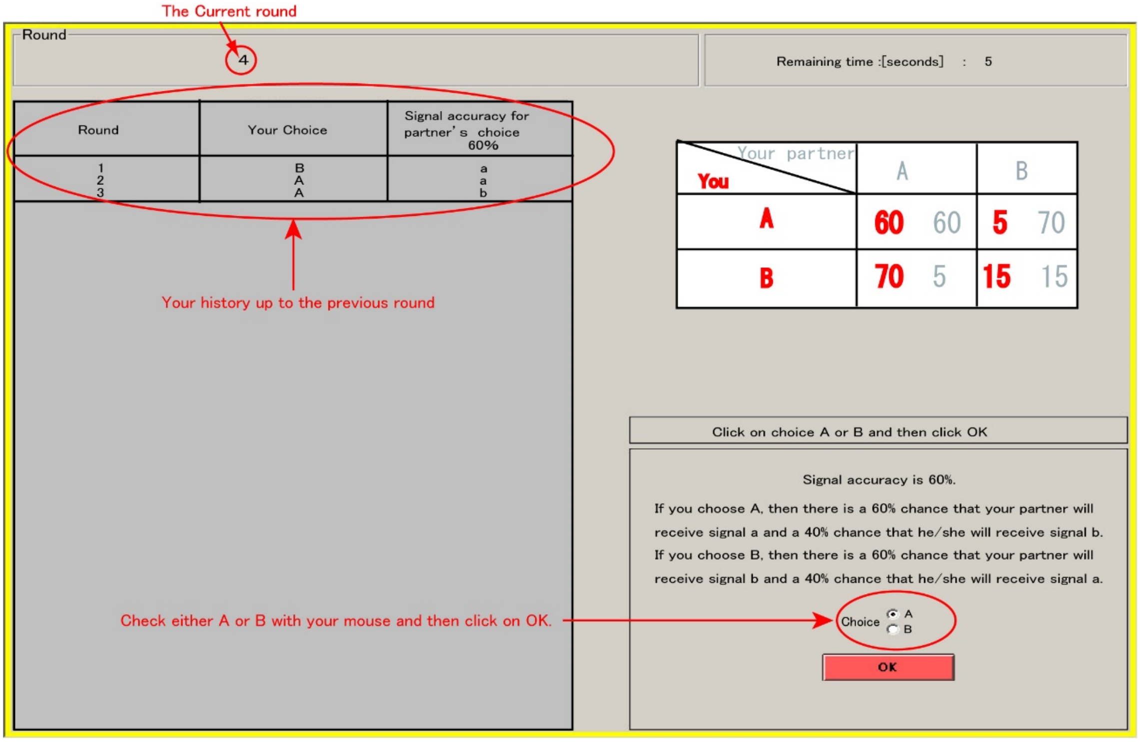
Screen 7: The Results Screen



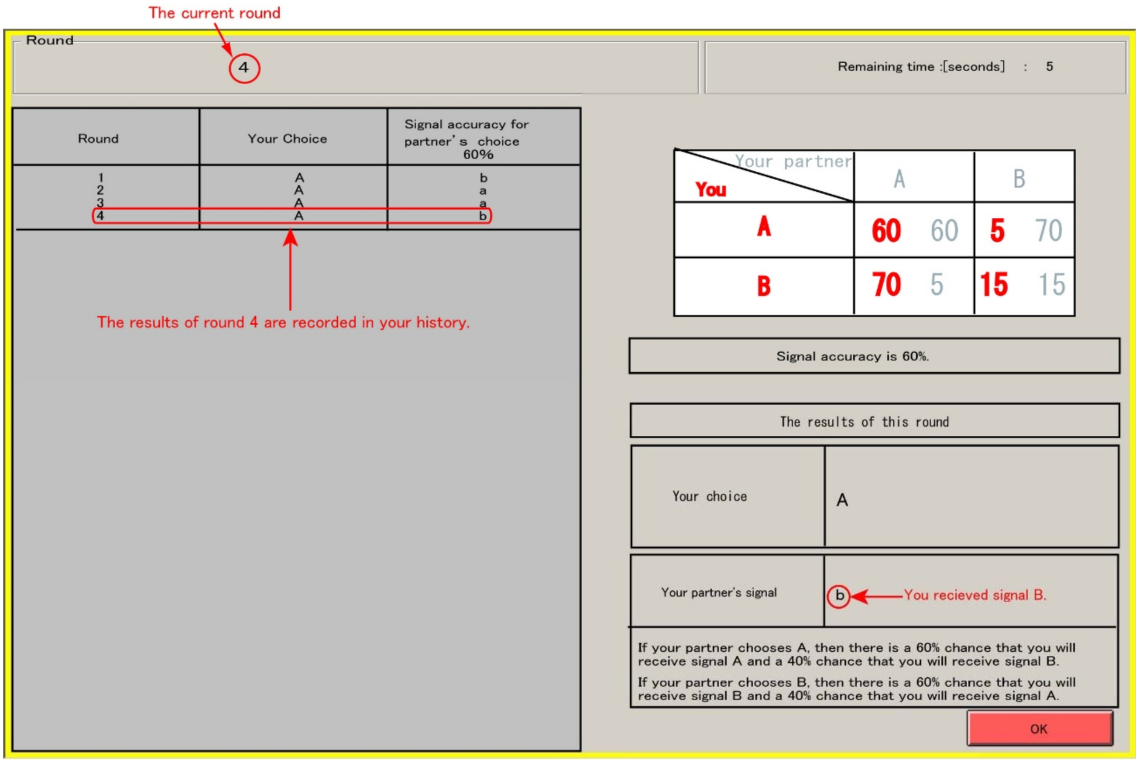
Screen 8: Your Selection Screen



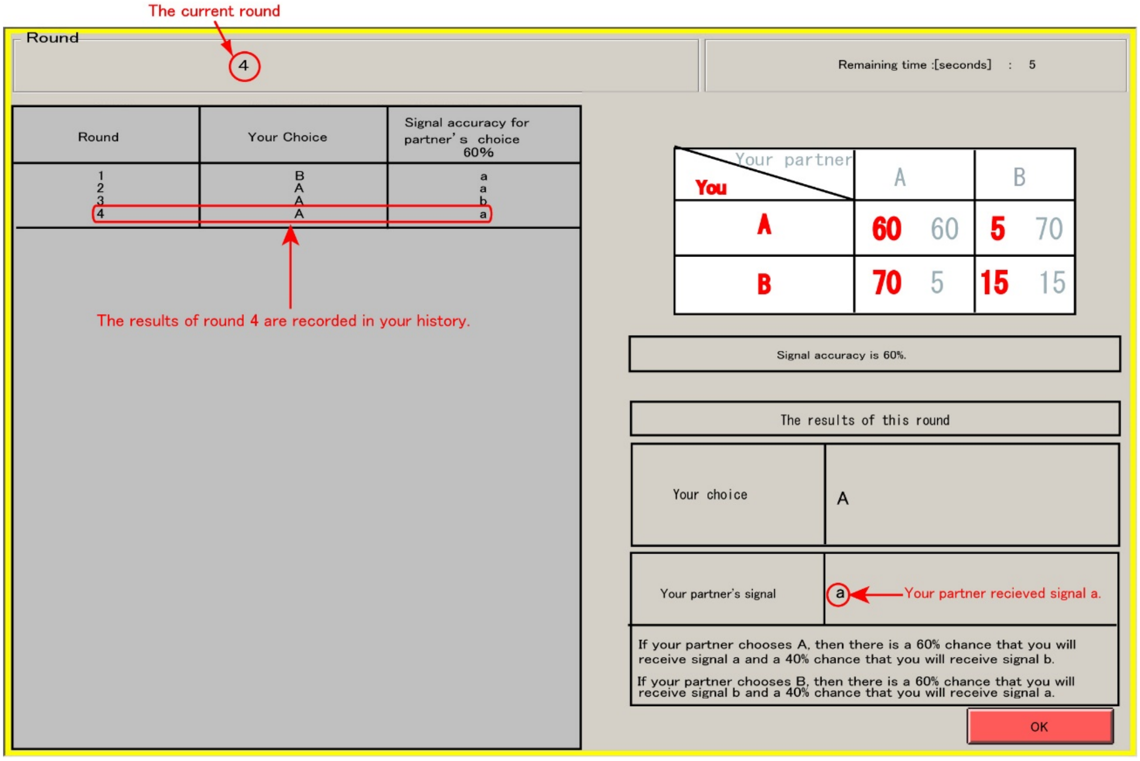
Screen 9: The Selection Screen of Your Partner



Screen 10: Your Results Screen



Screen 11: The Results Screen of Your Partner



Appendix I:
Experimental Instruction
(August 2018, Translation from Japanese into English)

1. Experimental Instruction

Please make sure all the contents are in your envelope. The envelope should have the following items.

1. Pen – 1
2. Instruction – 1 copy
3. Printed computer screen images – 1 copy
4. Bank transfer form – 1 sheet
5. Scratch paper – 1 sheet

If you are missing any item, please raise your hand quietly. We will collect the items at the end of all the experiments, except for the scratch paper, which you can keep.

Please look at the instructions (this material). You will be asked to make decisions at a computer terminal. You will earn "points" according to your performance in the experiments. The points will be converted into monetary rewards at the exchange rate of 0.9 yen per point, which will be paid in addition to the participation fee (1,500 yen). The total amount of money you will receive from the experiments is

the number of points earned × 0.9 yen + participation fee of 1,500 yen.

Any communication with other participants (i.e., conversation or exchange of signals) is not allowed during the experiments; if you violate this rule, you may be asked to leave the experiments. Furthermore, you are not allowed to leave in the middle of the experiments unless an experimenter allows or asks you to do so. Please turn off your cell phones during the experiments.

Outline of Experiments

We will conduct five experiments. The five experiments are independent of each other; the records of one experiment are not transferred to the other experiments. The experiments are conducted via a computer network. You are asked to make decisions at a computer terminal and interact with other participants through the computer network.

All the participants will be divided into pairs in each experiment. The pairs are selected randomly by the computer.

Each experiment consists of several rounds (i.e., Rounds 1, 2, 3, etc.). Later, we will explain the rule that decides the number of rounds conducted in each experiment. In each round, you are asked to choose one of two alternatives, which will also be explained below.

Please raise your hand quietly if you have any questions.

Decision Making

You will be asked to choose either A or B in each round. Your partner will also be asked to choose either A or B. Please look at the table.

You \ Your partner	A	B
A	60 60	5 70
B	70 5	15 15

The table summarizes the points you and your partner earn according to the combination of choices made by the two players. The characters in the left column marked in red indicate your choice, which is either A or B. The characters in the top row marked in light blue indicate the choice of your partner, which is also either A or B. In each cell, the numbers in red on the left side indicate the points you earn, and the numbers in light blue on the right side indicate the points your partner earns.

If both you and your partner select A,

both you and your partner earn 60 points.

If you select A and your partner selects B,

you earn 5 points, and your partner earns 70 points.

If you select B and your partner selects A,

you earn 70 points, and your partner earns 5 points.

If both you and your partner select B,

both you and your partner earn 15 points.

Please look at the table carefully and ensure that you understand how the points will be awarded to you and your partner according to the choices made by the two players. Your earnings depend not only on your choice but also on the choice of your partner. Similarly, your partner's earnings depend on your choice as well as her own.

Please raise your hand quietly if you have any questions.

The five experiments follow identical rules and will be conducted consecutively.

Observable Information

You are not allowed to observe whether your partner selected A or B directly. However, you will receive signal a or signal b, which has information about your partner's choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to you.

If your partner selects A, you will receive

signal a with a 90% chance and signal b with a 10% chance.

If your partner selects B, you will receive

signal b with a 90% chance and signal a with a 10% chance.

In the same way, your partner will not know whether you have selected A or B. However, your partner will receive signal a or signal b, which has information about your choice. According to the following rules, the computer determines stochastically whether signal a or signal b appears to your partner.

If you select A, your partner will receive

signal a with a 90% chance and signal b with a 10% chance.

If you select B, your partner will receive

signal b with a 90% chance and signal a with a 10% chance.

The signal you receive and the signal your partner receives are decided independently and have no correlation. Furthermore, the computer determines the signals independently in each round.

We refer to the stochastic rules for this signal-generating process as **signaling with 90% accuracy.**

Please raise your hand quietly if you have any questions.

Number of Rounds

The number of rounds in each experiment will be determined randomly. At the end of each round, the computer will randomly select a number from 1 to 30 without replacement, so there is a 1/30 chance of any number being selected by the computer. The number selected by the computer is applied uniformly to all participants.

The experiment will be terminated when the number 30 is selected by chance.

The experiment will continue if any number other than 30 is selected. However, you will notice only that a number other than 30 is selected, instead of seeing the specific number selected by the computer. Then, you will move on to the next round and will be asked to make a decision faced with the same partner.

When Experiment 1 is terminated, you proceed to Experiment 2, and you will be randomly paired with a new partner. When Experiment 2 is terminated, you proceed to Experiment 3, and again, you will be randomly paired with a new partner. When Experiment 3 is terminated, you proceed to Experiment 4, and again, you will be randomly paired with a new partner. When Experiment 4 is terminated, you proceed to Experiment 5, and again, you will be randomly paired with a new partner.

Please raise your hand quietly if you have any questions.

Description of Screens and Operations for Computers

Please look at the booklet with printed computer screen images.

Please look at Screen 1 and Screen 2. Screen 1 displays the screen that will be presented to you during the decision phases. Screen 2 is the screen that will be presented to your partner during the decision phases. Please look at the top left portion of each screen, which indicates that the current round is Round 4. The left portion of Screen 1 displays the information available to you up to the round. The left portion of Screen 2 presented to your partner displays the information available to her up to the round.

You are asked to click with the mouse to select either “A” or “B” in the bottom right portion of the screen. Then, the selection will be confirmed by clicking the “OK” button right below the alternatives.

Next, please look at Screen 3 and Screen 4. Screen 3 presents the results to you. Screen 4 presents the results to your partner. The screens display the situation in which, in Round 4, both you and your partner chose A. Screen 3 shows you that, in Round 4, “your partner’s signal (accuracy: 90%) is b,” indicating to you that the signal you observe about the partner’s choice is “b.” On the other hand, Screen 4 shows your partner that, in Round 4, “your partner’s signal (accuracy: 90%) is a,” indicating to your partner that the signal your partner observes about your choice is “a.” Recall that your partner will observe signal a with a probability of 90% and will observe signal b with a probability of 10% when you choose “A.”

Then, we move on to the lottery screens. Please turn the page and look at Screen 5 and Screen 6, which display the lottery. Any number from 1 to 30 will be randomly selected with an identical probability of occurrence, which is 1/30. Then, a part of the cells turns green according to the number selected. If the number 30 is selected, the cell numbered 30 turns green and the message below explains that the current experiment is terminated.

Otherwise, Screen 5 is shown, in which all the cells numbered 1 to 29 turn green at once (you do not know which number is selected specifically), and the message below explains that the experiment continues with the same partner. Screen 6 is presented when the number 30 is selected, and the cell numbered 30 turns green, indicating that the current experiment is terminated in that round. Again, please make sure that the experiment is terminated when the cell numbered 30 turns green.

Please raise your hand quietly if you have any questions.

Now, all the processes of the experiments have been completed, and all the points awarded to everyone recorded on the computer.

Please answer the questionnaire that will be distributed now.

Take the bank transfer form out of the envelope and fill it out accurately; otherwise, we will not be able to process the payment correctly for you.
Please raise your hand quietly if you have any questions.

Please make sure that you fill out the questionnaire and the bank transfer form correctly.
Please raise your hand quietly if you have any questions.

Please put all the documents in the envelope. Please leave the pen and ink pad on the desk. Make sure you take all your belongings with you when you leave.

Please do not disclose any details regarding the experiments to anyone. Thank you very much for your participation. Please follow the instructions of the experimenters to leave the room.