CIRJE-F-1125

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

Yutaka Kayaba
The University of Tokyo
Hitoshi Matsushima
The University of Tokyo
Tomohisa Toyama
International Christian University

September 2019

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: September 26, 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 the 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. Showing inconsistent behavior, subjects with low accuracy do not tend to retaliate more than those with high accuracy. Furthermore, subjects with low accuracy tend to retaliate considerably with lesser strength than that predicted by the equilibrium theory, while subjects with high accuracy tend to retaliate with more strength than that predicted by the equilibrium theory, or with strength almost equivalent to it.

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); Matsushima, Tanaka, and Toyama (2013); Kayaba, Matsushima, and Toyama (2018) are earlier versions of this paper. Research for this study was financially supported by a grant-in-aid for scientific research (KAKENHI, grant numbers: 16H02009, 18K12742) 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, the advisory editor, and the anonymous reviewers for their helpful comments and suggestions. We are grateful to Hide Ichimura and the participants of the 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

Long-run strategic interactions facilitate collusion among players whose interests conflict with that of each other. The premise is that each player observes the actions that opponents have selected previously and retains this information. However, even if the monitoring of opponents' actions is *imperfect* (i.e., each player cannot directly observe the opponents' action choices but can observe only the informative signals they generate), theoretical studies have shown that patient players can still employ, to a greater or lesser degree, cooperative strategies as an equilibrium. To be more precise, the folk theorem 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 can be applied to a wide range of strategic conflicts.

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

Our experimental setup is imperfect monitoring. Each player cannot directly observe his or her opponent's action choice; instead, he or she observes a signal, which is either *good* or *bad*. The good (bad) signal is more likely to occur when the opponent selects a cooperative (defective) action rather than a defective (cooperative) action. The probability of a player observing the good (bad) signal when the opponent selects the cooperative (defective) action is referred to as monitoring accuracy; it is denoted by $p \in (1/2, 1)$. The study experimentally controls the levels of monitoring accuracy as treatments (high accuracy $p = 0.9$ and low accuracy $p = 0.7$). Specifically, the monitoring technology is *private* in that a player does not receive any information about what the opponent observes about his or her choices (i.e., the signals are observable only by the receivers).

To examine the strategies used by the subjects, 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, long-memory (lenience), and long-term punishment strategies, all of which comprise a significant proportion of the strategies found in existing studies of experimental repeated games. SFEM then allows us to estimate the frequencies of each strategy, wherein the heterogeneity of the strategies followed by the subjects is treated explicitly. Existing experimental studies use 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 his or her opponent's action by making a cooperative (defective) action choice whenever he or she observes a good (bad) signal. G-TFT is defined as the stochastic extension of TFT, according to which, the probability of a player making a cooperative (defective) action choice is higher when he or she observes a good (bad) signal than when he or she observes a bad (good) signal.

TFT is considered 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 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 they fall 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 is reasonable to expect human beings and animals to conduct such random experimentations as g-TFT implies. Indeed, evolutionary biology finds that animals maintain peaceful coexistence, instead of weak costs, by adopting g-TFT (e.g., Molander, 1985; Nowak and Sigmund, 1992). In human societies, g-TFT provides an opportunity to avoid the crisis of

[5] Axelrod (1984), and Wu and Axelrod (1995) showed that TFT and g-TFT are 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 contrary, some game theorists criticize the insufficiency of Axelrod's evolutionary simulations (e.g., Binmore, 1994, Chapter 3).

nuclear war and build peaceful relationships among countries. 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 subjects playing g-TFT strategies, as well as those playing other strategies that are common in existing studies of experimental repeated games. We also include various long-memory (lenient) strategies and their stochastic versions in the list.

Our estimates indicate that a significant proportion (approximately 60–80%) of our subjects follow g-TFT strategies, albeit heterogeneous ones. Although existing empirical or theoretical studies emphasize grim, long-memory (lenient), and long-term punishing strategies, our experimental results demonstrate that the empirical importance of g-TFT is supported experimentally as well as theoretically as 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, retaliation intensities are common across all g-TFT equilibria depending on the level of monitoring accuracy. As the level of monitoring accuracy decreases, common retaliation intensity increases. This property plays a central role in improving monitoring technology and effectively saving welfare loss caused by the monitoring imperfection.

However, the retaliation intensities observed in our experimental data are contrary to the predictions of the above-mentioned equilibrium theory; our subjects with low accuracy do not retaliate more than those with high accuracy. To be precise, the retaliation intensity with low accuracy is slightly lesser than, or at the most equivalent to, the retaliation intensity with high accuracy. Furthermore, subjects with low accuracy tend to retaliate with lesser strength than that expected by the equilibria, while subjects with high accuracy tend to retaliate with more than the expected strength, or with the almost equivalent strength, implied by the equilibria. Hence, when monitoring is inaccurate, the expected payoff from cooperation tends to be considerably less than that from defection and when monitoring is accurate, the expected payoff to an individual subject from

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

cooperation tends to be greater 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 results for individual strategies. Section 9 is the conclusion.

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 used the self-generative nature of perfect equilibria explored by Abreu (1988) and Abreu, Pearce, and Stacchetti (1990), which, however, crucially relied on the publicity of signal observations.

In studying imperfect private monitoring, Ely and Välimäki (2002) and Piccione (2002) explored 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 which monitoring is private and far from perfect.

The literature of experimental studies on repeated games has examined the

[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.

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 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 from the SFEM list. The inclusion of such stochastic action choices is scant in the literature on 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 employ lenient (i.e., long-memory) strategies.[8]

By contrast, we rigorously include many variants of g-TFT in our SFEM list. We also include various long-memory (lenient) 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 occurrences 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

The 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, is 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 $u_i(s)$ denotes the payoff for player i induced by action profile $s \in S$.

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

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 $f(\cdot | s): \Omega \to R_+$, where $\sum_{\omega \in \Omega} f(\omega | s) = 1$ for all $s \in S$. We assume *full support* in that $f(\omega | s) > 0$ for all $\omega \in \Omega$ and $s \in S$.

Let $f_i(\omega_i | s) \equiv \sum_{\omega_j \in \Omega_j} f(\omega | s)$, and we assume that $f_i(\omega_i | s)$ is independent of s_j. Hence, we denote $f_i(\omega_i | s_i)$ instead of $f_i(\omega_i | s)$. We use $\omega_i \in \Omega_i$ to denote *the signal for player* $i's$ *action choice.* Player $i's$ action choice s_i influences the occurrence of the signal for his or 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 $t \in \{1,2,...\}$, 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 $u(s(t)) = (u_1(s(t)), u_2(s(t))) \in R^2$, in which the action profile selected in round t is denoted by $s(t) = (s_1(t), s_2(t)) \in S$. Instead, player i observes the signal for opponent $j's$ action choice $\omega_j(t) \in \Omega_j$ by 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 is unable to know the type of signal his or her opponent receives about his or 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: $S_1 = S_2 = \{C, D\}$; $u_1(C,C) = u_2(C,C) = 1$; $u_1(D,D) = u_2(D,D) = 0$; $u_1(C,D) = u_2(D,C) = -g$; and $u_1(D,C) = u_2(C,D) = 1 + g$,

[9] Our specification of the 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.

where it is assumed that $g > 0$. 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 a unique Nash equilibrium.

We specify $\Omega_i = \{c,d\}$, $f_i(c \mid C) = f_i(d \mid D) = p$, 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 the value of p, 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 rather than D.

Let $h(t) = (s(\tau), \omega(\tau))_{\tau=1}^{t}$ denote the *history up to round* t. $H = \{h(t) \mid t = 0,1,...\}$ 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, he or 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 the 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 $p \in (0,1)$ and is defined as

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

where $E[\cdot \mid \sigma, p]$ 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

$$U_i(\sigma; p) \geq U_i(\sigma_i', \sigma_j; p) \quad \text{for all} \quad i \in \{1,2\} \quad \text{and all} \quad \sigma_i' \in \Sigma_i \text{.}$$[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 a *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 $h(t-1) \in H$,

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

and

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

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 he or 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 equilibrium in the repeated game with accuracy p.

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

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 the 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 seen to *decrease* in p; this implies that the less accurate the monitoring technology, the more severely the 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 his or 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 a bad signal even if the player has actually made the cooperative action choice. This inevitably leads to welfare loss, because the opponent might retaliate against the player even if he or she has made the cooperative action choice. In such as 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 eight sessions of computer-based laboratory experiments at the Center for Advanced Research for Finance, University of Tokyo, in October 2018 and June 2019.[11] We recruited 224 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 (0.9 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 a variable coefficient

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

45 and a 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 made 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 where monitoring technology is poorer. The chance of the signals observed by the player matching the opponent’s action choices is only 70% ($p = 0.7$). We refer to this treatment as the “low accuracy treatment.” The 224 subjects were divided into two teams of 112 each, with one team assigned to the high accuracy treatment and the other to the low accuracy treatment, respectively.

After playing a short, repeated game of two rounds as practice, each subject plays five repeated games in the assigned treatment. At the start of each repeated game, subjects were randomly paired, and the pairs remained unchanged until the end of the repeated game. With respect to the determination of the final round in each repeated game, 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 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 ($p = 0.7$).

Employing the continuation probability mentioned above, we randomly generate four sequences of five numbers, and each sequence is used as the number of rounds for five repeated games of a session once in both treatments; each treatment includes four identical sequences of repeated games. This procedure for matching the sequences of repeated games across the two treatments erases the possible heterogeneities in the learning effects owing to the randomness of the number of experienced rounds across treatments. Table 2 displays the summary of sequences of repeated games, their treatment, and the number of subjects.

[TABLE 2 HERE]

Our subjects were not informed in advance about which was the final round in each repeated game. To help our subjects understand that the probability of termination used for generating the sequences of repeated games is 1/30, we presented 30 cells (numbered 1 to 30) on the computer screen at the end of each round. The 30th cell in the computer screen turned green when the repeated game was terminated. Otherwise, all the cells numbered 1 to 29 turned green simultaneously and the repeated game continued. The screen is demonstrated in the Online Appendix C.

Using printed experimental instructions, each subject was informed of the rules of the game and the ways in which the game would proceed. The instructions were explained using a recorded voice. Using the computer screen during the experiments, our subjects could review the structure of the game and the history up to the latest round, the history of one's own actions, and the signals of the opponent's actions. See Online Appendix G for the experimental instructions and the images on the computer screen, which have been translated into English from the original Japanese.

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. In each treatment of monitoring accuracy, a group of 112 subjects made 16,968 decisions. The overall frequency of cooperative choices (i.e., the cooperation rate) is 0.845 in the high accuracy treatment, and 0.598 in the low accuracy treatment. Statistically, the former frequency is significantly larger than the latter ($p < 0.001$). A first look at the cooperation rates suggests that our subjects cooperate more as the monitoring technology improves.

[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 0.905 in the high accuracy treatment and 0.748 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 not only in the high accuracy treatment but even in the low accuracy treatment. However, their initial motivation for cooperation diminishes slightly with an increase in the noise in the signal.

6.2. Signal-Contingent Cooperation Rates

Table 4 also presents the 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 0.924. 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.[12] The mean of individual-level means in the high accuracy treatment is 0.918, which is 0.006 less than the simple mean of choices, implying that there could be over-representation. However, both measures are consistently high, touching 0.9, thereby indicating that our subjects are quite cooperative when they observe a good signal.

Even in the low accuracy treatment, the cooperation rate after observing a good signal is high, although it is not as high as in the high accuracy treatment. The simple

[12] 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.

mean of the cooperative choices in the low accuracy treatment is 0.726, and the mean of individual-level means is 0.706. As in the case of the round 1 cooperation rate, to some extent, our subjects are reluctant to cooperate even after observing good signals in the low accuracy treatment. A 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 0.560, and the mean of individual-level means is 0.621. Both measures are considerably smaller than that after observing a good signal in the treatment (Table 5). Bad signals tend to make our subjects consider more of defection.

The tendency of our subjects towards defection upon getting a bad signal is also observed in the low accuracy treatment. The simple mean of cooperative actions across all choices is 0.437 and the mean of individual-level means is 0.460, both of which are consistently smaller than the means in the high accuracy treatment ($p = 0.002$ for the simple mean, and $p < 0.001$ for the individual-level mean). Observing a bad signal, our subjects tend to defect more in the low accuracy treatment than in the high accuracy treatment. Again, both measures are considerably smaller than that after observing a good signal in the treatment (Table 5).

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. These findings imply that the strategies our subjects employ differ with the change in signal quality.

RESULT 1-a: Our subjects in the high accuracy treatment tend to cooperate more than those in the low accuracy treatment, and thus the strategies they employ differ according to the monitoring 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 0.365 in the simple mean of all choices and 0.297 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. However, both measures are larger and statistically significant than the level implied by the g-TFT equilibria ($w(0.9)=0.235$, $p<0.001$ for the simple mean, and $p=0.003$ for the individual-level mean), although the difference between the theoretical level and the individual-level mean is only 0.062. Thus, empirically, our subjects tend to punish partners slightly more than necessary, or almost just enough, to incentivize them to collude in the high accuracy treatment.

This deviation from the equilibria can be perceived more in the low accuracy treatment. The retaliation intensity $r(c;0.7)-r(d;0.7)$ is 0.289 in the simple mean of all choices and 0.247 in the mean of individual-level means. Both measures consistently and with statistical significance differ from 0 ($p<0.001$ for both), which demonstrates that our subjects use signal-contingent information even with poorer monitoring technology. However, unlike the case of the high accuracy treatment, the retaliation intensity in the low accuracy treatment is considerably lower than the level implied by the g-TFT equilibria ($w(0.7)=0.47$) for both measures ($p<0.001$ for both). Although our subjects retaliate according to the signals even in the low accuracy treatment, the strength of the retaliation is considerably below the theoretical level, allowing the opponents to defect permanently to pursue larger payoffs.[13]

---

[13] There might be concerns 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.

Not only do the retaliation intensities mismatch with the value implied by the g-TFT equilibria, but the players do not retaliate strongly with low accuracy than with high accuracy as predicted by the g-TFT equilibria. The direct comparison of the observed retaliation intensities across the two treatments (presented in Table 5) indicate that the retaliation intensity in the low accuracy treatment is not larger than the retaliation intensity in the high accuracy treatment, while the former is larger than the latter by a value of 0.235 as predicted by the equilibrium theory ($w(0.7) - w(0.9) = 0.47 - 0.235 = 0.235$). Rather, the observed retaliation intensity in the low accuracy treatment is slightly smaller than that in the high accuracy treatment as evaluated from the point estimates (statistically significant or marginally significant, $p = 0.023$ for the simple mean, and $p = 0.079$ for the individual-level mean).

RESULT 2-a: In the low accuracy treatment, our subjects tend to retaliate lesser than the level implied by the g-TFT equilibria, while they tend to retaliate at a higher level than, or at the same level, as implied by the g-TFT equilibria in the high accuracy treatment. Inconsistent with the theoretical predictions, the retaliation intensity does not increase when the monitoring technology becomes poorer.

6.4. 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, we perform reduced-form linear regression analyses. The results indicate that, despite some learning effects observed in the early stage of the experiments, the final three repeated games in both treatments have almost none or insignificantly small learning effects uniformly among overall cooperation rates, signal contingent-cooperation rates, and retaliation intensities. A detailed discussion on this is given in Online Appendix D.

However, our analysis in Section 8 indicates that only a small number of our subjects adopted long-term punishing strategies, and, hence, the effect of long-term punishments is minimal.

7. Estimation of Individual Strategies—Methodology

In this section and the following one, we present direct estimations of individual strategies of our subjects. Given the recent consensus in the literature of experimental repeated games that a substantial amount of 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 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 validity of this method is verified in the Monte Carlo simulations in Fudenberg, Rand, and Dreber (2012). We also perform simulation exercises to examine the validity of the SFEM that includes a larger set of g-TFT strategies in the strategy list with the sample size of the current study; the results are reported in Online Appendix E.

The underlying assumption for SFEM is that each subject continues to employ a specific strategy across all repeated games in each treatment. As discussed in Online Appendix D, the final three repeated games in both treatments include almost none or insignificant learning effects. This implies that there is a lesser possibility of the subjects systematically changing their strategies in the final three repeated games. Thus, we employ the data in the final three repeated games for the SFEM for both treatments. However, we also perform estimations employing the final four repeated games as robustness checks and find few changes in the estimates in each treatment; this would not be the case if our subjects systematically changed their strategies. The summary of the estimation results is documented in Online Appendix F.

[TABLE 6 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 6 displays the list of strategies in our SFEM. The list includes TFT, TF2T, TF3T, 2TFT, 2TF2T, Grim (trigger strategy),[14] Grim-2, Grim-3, always cooperate (ALL-C), and always defect (ALL-D), all of which are listed in the literature of infinitely repeated games in imperfect monitoring (e.g., Fudenberg, Rand, and Dreber, 2012).[15],[16]

Among these, ALL-D is a non-cooperative strategy, while the others (TFT, TF2T, TF3T, 2TFT, 2TF2T, Grim, Grim-2, Grim-3, and ALL-C) are cooperative strategies that involve cooperation at the start of each repeated game and continue to involve cooperation unless the belief sets in that 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.

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

[14] 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 he or she observes a bad signal or has played defection in the previous round. Note that he or she could mistakenly play defection before the "trigger" is pulled because implementation errors of action choices are allowed in the SFEM framework.

[15] 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.

[16] The literature has often added to the strategy set Perfect tit-for-tat (P-TFT), wherein cooperation is chosen if 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.

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 $r(d)$.[17] 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.[18]

We also include a family of g-TF2T (g-TF2T-r) as the generous variants of TF2T; these allow stochastic punishments if two consecutive bad signals occur (play cooperation

---

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

[18] 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.

with probability after observing two consecutive bad signals).[19]

8. Estimation of Individual Strategies–Results

8.1. Cooperative and Non-Cooperative Strategies

[Table 7 HERE]

[Table 8 HERE]

Table 7 presents the estimates for the frequencies of strategies our subjects follow, and Table 8 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 98.8%, while that of non-cooperative strategies is 1.2%; thus, the share of cooperative strategies exceeds that of non-cooperative strategies. Statistically, the latter is significantly smaller than the former ($p < 0.001$). Although there are considerable heterogeneities in the strategies that our subjects follow, as Table 7 shows, most of our subjects adopt cooperative strategies in the high accuracy treatment.

In the low accuracy treatment, the share of cooperative strategies is 88.2%, while that of non-cooperative strategies is 11.8%; again, the share of non-cooperative strategies exceeds that of cooperative strategies. Statistically, the latter is significantly smaller than the former ($p < 0.001$). Many of our subjects follow cooperative strategies even when the monitoring technology is considerably poor, although the share of non-cooperative strategies in the low accuracy treatment is significantly larger than in the high accuracy treatment ($p = 0.007$).

---

[19] 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.

RESULT 1-b: Many of our subjects follow cooperative strategies in both the high and low accuracy treatments, although the frequency of cooperative strategies is larger in the high accuracy treatment than in the low accuracy treatment.

Remark: Fudenberg, Rand, and Dreber (2012) reported in their imperfect public monitoring experiment that frequencies of cooperative strategies drop as the level of noise is increased from 1/16 to 1/8. Our study demonstrates that, even with the more drastic change in signal noise from 1/10 to 3/10 in imperfect private monitoring, the poorer signal quality does not drastically discourage our subjects from adopting cooperative strategies.

8.2. Proportion of g-TFT Strategies

We now examine the share of g-TFT. Our SEFM estimates in Tables 7 and 8 indicate that there is a substantial proportion of the g-TFT family in our imperfect private monitoring. In the high accuracy treatment, independently, g-TFT-0.875-0.5 has the highest share among the various strategies (13.4%), followed by g-TFT-1-0.75 with an almost identical share (13.1%). 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 52.6%. The extended family of g-TFT, which includes the signal non-contingent, primitive variants of g-TFT, has a 62.5% share. As shown by these large numbers, a considerable proportion 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 58.5%, while the extended family with signal non-contingent variants comprises 78.1% of the strategies.

Regardless of the treatment, we find a substantial proportion of our subjects following strategies in g-TFT family. This indicates that our subjects' decisions on retaliation largely depend on a single occurrence of a bad signal (c.f. long-memory strategies).[20]

---

[20] The result that our SFEM estimates find many of our subjects playing g-TFT in our imperfect

RESULT 3: Our SFEM estimates indicate that the g-TFT family comprises a substantial portion of the strategies our subjects follow in both treatments.

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-0.625-0.375, g-TFT-0.5-0.25, g-TFT-0.375-0.125, and g-TFT-0.25-0 have approximately the same retaliation intensity as implied by the g-TFT equilibria ($w(0.9) = 0.235$) in our list. However, our SFEM estimates in Tables 7 and 8 indicate that the joint share of these strategies is only 13.1%, though it is statistically significantly different from 0 at the 5% significance level ($p = 0.029$); a very small fraction of our subjects follow g-TFT equilibria retaliation intensity in the high accuracy treatment. This finding also holds for the low accuracy treatment in which g-TFT-1-0.5, g-TFT-0.875-0.375, g-TFT-0.75-0.25, g-TFT-0.625-0.125, and g-TFT-0.5-0 have approximately the same retaliation intensity as implied by the g-TFT equilibria ($w(0.7) = 0.47$). However, the joint share of these strategies in the low accuracy treatment is only 4.6%, which is not statistically significantly different from 0 ($p = 0.321$). Despite many subjects following one of the g-TFT strategies in both treatments, very few of them follow the retaliation intensities implied in the g-TFT equilibria.

We further address how they tend to deviate from the theoretical prediction in terms of proportions of strategies. Our SFEM estimates in Tables 7 and 8 indicate that the group of stronger retaliation variants of g-TFT, that is, g-TFT strategies with retaliation

---

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, we corroborate Fudenberg, Rand, and Dreber (2012) by finding considerable proportions of our subjects following long-memory strategies in both treatments.

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 29.1% in the high accuracy treatment (significant, $p < 0.001$). Since approximately 60% of our subjects follow g-TFT in the high accuracy treatment, the results indicate that roughly half of them retaliate more strongly than the g-TFT equilibria require. Moreover, 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, reaches a somewhat equivalent share, which is 20.4% (significant, $p = 0.002$), as the share does not statistically significantly differ from that of the group of stronger retaliation variants ($p = 0.327$). In the high accuracy treatment, the deviation from equilibria is roughly unsystematic, or slightly towards stronger retaliation as evaluated from the point estimates.

The deviation from the equilibria is systematic in the low accuracy treatment. Our SFEM estimates indicate that the group of weaker retaliation variants of g-TFT, that is, g-TFT-1-0.875, g-TFT-1-0.75, g-TFT-1-0.625, g-TFT-0.875-0.75, g-TFT-0.875-0.625, g-TFT-0.875-0.5, g-TFT-0.75-0.625, g-TFT-0.75-0.5, g-TFT-0.75-0.375, g-TFT-0.625-0.5, g-TFT-0.625-0.375, g-TFT-0.625-0.25, g-TFT-0.5-0.375, g-TFT-0.5-0.25, g-TFT-0.5-0.125, g-TFT-0.375-0.25, g-TFT-0.375-0.125, g-TFT-0.375-0, g-TFT-0.25-0.125, g-TFT-0.25-0, ALL-C, ALL-D, and Random-r, jointly comprise 64.2% (significant, $p < 0.001$), while strong retaliation variants of g-TFT comprise only 9.3% (significant at the 5% level, $p = 0.017$); indeed, the group of weaker retaliation is significantly larger than the group of stronger retaliation ($p < 0.001$). Unlike the high accuracy treatment, the deviation from equilibria is systematic towards weaker retaliation in the low accuracy treatment.

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 of all players. 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 0.312 (s.e. 0.041); this is marginally significant and larger than the value predicted by the g-TFT equilibria ($w(0.9) = 0.235$, $p = 0.058$). In the low accuracy treatment, the mean retaliation intensity is 0.274 (s.e. 0.029) ; this is considerably smaller than the value implied by the g-TFT equilibria ($w(0.7) = 0.47$, $p < 0.001$). A comparison of the two indicates that the conditional mean of retaliation intensity in the low accuracy treatment is not larger than that in the high accuracy treatment by the degree that the g-TFT equilibria imply ($w(0.7) - w(0.9) = 0.47 - 0.235 = 0.235$, $p < 0.001$). Rather, the retaliation intensity in the low accuracy treatment is smaller than in the high accuracy treatment as evaluated from the point estimates, though the difference is not statistically significant ($p = 0.452$). Hence, even if we consider only g-TFT players, the behavior deviates from the theoretical predictions; this result echoes the findings in Section 6.

RESULT 2-b: Our SFEM estimates indicate that, only a small number of our subjects follow the retaliation intensities implied by the g-TFT equilibria in both treatments. Rather, the share of weaker retaliation variants of g-TFT considerably outweighs that of stronger retaliation variants in the low accuracy treatment, while in the high accuracy treatment, the share of stronger retaliation variants of g-TFT is equivalent to that of weaker retaliation variants of g-TFT with some possibility that the former might be slightly larger than the latter. The mean retaliation intensity among g-TFT players is considerably smaller than the level implied by the equilibria in the low accuracy treatment, and is marginally larger than the level implied by the g-TFT equilibria in the high accuracy treatment. Inconsistent with the predictions of g-TFT equilibria, the retaliation intensity of g-TFT players in the low accuracy treatment does not tend to be larger than in the high accuracy treatment.

8.4. Further Results

8.4.1. Long-Term Punishment

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. Here we address the concern that seemingly weak retaliation intensity might spuriously arise when many subjects employ long-term punishing strategies.

Tables 7 and 8 indicate that, 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 only 12.8%; the share is statistically only marginally significant ($p = 0.081$). Likewise, in the high accuracy treatment, the share of long-term punishing strategies is 14.2% (significant, $p < 0.001$).

Since only a small number of the subjects employ long-term punishing strategies in both treatments, their effects are minimal. Indeed, the mean retaliation intensity among g-TFT players in the low accuracy treatment reported in this section, that is 0.274, is somewhat larger than the value reported for the mean retaliation intensity in Section 6, that is 0.247 (individual-level mean). However, as shown above, the value is still considerably below the level implied by the g-TFT equilibria ($w(0.7) = 0.47$).

RESULT 2-c: The shares of strategies with long-term punishments are only marginal or at most small in both treatments.

8.4.2. Long-Memory Strategy

Given the theories regarding review strategies (e.g., Radner, 1986; Matsushima, 2004; Sugaya, 2012), players might rely on signals in longer histories to compensate for the informational disadvantages of poorer monitoring technologies. Therefore, it is interesting to examine whether our subjects tend to play long-memory strategies more in the low accuracy treatment than in the high accuracy treatment as the theories of review strategies suggest.

Tables 7 and 8 indicate that, in the low accuracy treatment, the joint share of long-memory strategies, that is, the family of g-TF2T including TF2T, TF3T, 2TF2T, Grim-2, and Grim-3, is 21.0% (marginally significant, $p = 0.078$). Individually, Grim-3 has the largest proportion, which is 7.6% (significant at the 5% level, $p = 0.021$). In the high accuracy treatment, their joint share is 37.5% (significant, $p < 0.001$). Individually, g-

TF2T-0.5, which is a hybrid of generous and long-memory strategies, has the highest frequency, which is 15.6% (significant at the 5% level, $p = 0.042$). These results indicate that considerable proportions of our subjects employ long-memory strategies in both treatments; however, contrary to the theories of review strategies, the share does not increase in the poorer monitoring technology.

RESULTS 4: Our estimates indicate that a considerable number of subjects employ long-memory strategies, however there is no evidence of a larger share of the strategies in the low accuracy treatment as opposed to the high accuracy treatment.

Remark: These results corroborate previous findings in the literature; Fudenberg, Rand, and Dreber (2012) and Aoyagi, Bhaskar, and Fréchette (2019) find that a substantial number of players employ long-memory (lenient) strategies in imperfect monitoring. Moreover, similar to our case, Fudenberg, Rand, and Dreber (2012) find paradoxical results that the share of the long-memory players somewhat decreases as the monitoring technology becomes poorer. The results observed in this study echo these findings.

This study originally finds that, by adding g-TFT strategies extensively in the SFEM list, memory-one strategies (i.e., g-TFT strategies) are employed substantially, whose frequency is greater than, or at least equivalent to the frequency of long-memory (lenient) strategies; in the high accuracy treatment, the frequencies of g-TFT strategies excluding signal non-contingent strategies (i.e., non-trivial memory-one strategies) is 52.6%, which exceeds that of long-memory strategies (37.5%) as evaluated from the point estimates. However, the difference is not statistically significant ($p = 0.261$). Similarly, in the low accuracy treatment, the frequencies of g-TFT strategies excluding signal non-contingent strategies is 58.5%, which is larger than that of long-memory strategies (21.0%) as evaluated from the point estimates; however, the difference is only marginally significant ($p = 0.097$).

9. 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 attempt to investigate imperfect private monitoring.

Our results indicate that a significant proportion of the subjects employ 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, long-term punishment strategies, and their stochastic variants that frequently appear in existing experimental studies on repeated games. The proportion is somewhat greater than, or at least equivalent to the proportion of long-memory (lenient) strategies.

Although many subjects follow g-TFT strategies, their retaliating policies do not follow the predictions of the g-TFT equilibria. Inconsistent with the predictions, our subjects do not retaliate more with low accuracy than with high accuracy; their retaliation intensity in the low accuracy treatment is somewhat smaller than, or at most equivalent to that in the high accuracy treatment. Further, the subjects tend to retaliate considerably less than what the standard g-TFT equilibria predict in the low accuracy treatment, while they tend to retaliate somewhat more than, or almost equivalently to what is required by the equilibria in the high accuracy treatment. Hence, the expected payoff from cooperation tends to be considerably less than the expected payoff from defection when monitoring is inaccurate; however, the expected payoff to a subject from cooperation tends to be somewhat greater than that from defection when monitoring is accurate.

These 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 present to economists a new issue related to incentives that encompasses the motivations for retaliations beyond just maximizing pay-off.[21],[22]

Further, it is meaningful to attempt various experimental designs and examine subjects' choice of strategies more systematically. From this study and the previous experimental ones concerning imperfect monitoring, such as Fudenberg, Rand, and Dreber (2012), we learn that g-TFT and long-memory (lenient) strategies are remarkable in implicit collusion. However, grim-trigger and long-term punishing strategies are not so, even though they are prominent in theory and in real cartels (e.g., Igami and Sugaya, 2018). Further experimental investigation is needed to uncover various potential factors that could affect players' choices of strategies. Such factors might include pre-play communication and information transmission during the play; these are expected to make experimental environments resemble reality.

References

Abreu, D., 1988. On the Theory of Finitely Repeated Games with Discounting. Econometrica. 56, 383–396.

Abreu, D., Pearce, D., Stacchetti, E., 1990. Toward a Theory of Discounted Repeated Games with Imperfect Monitoring. Econometrica. 58, 1041–1063.

Aoyagi, M., Bhaskar, V., Fréchette, G., 2019. The Impact of Monitoring in Infinitely Repeated Games: Perfect, Public, and Private. American Economic Journal: Microeconomics. 11(1), 1-43.

Aoyagi, M., Fréchette, G. 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.

[21] The companion paper, Matsushima (2019) demonstrates a behavioral model of an infinitely repeated prisoners' dilemma.

[22] An earlier version of the paper, Kayaba, Matsushima, and Toyama (2018), analyze data from other experiments in which the experimental treatments are almost identical to the ones in the current paper, but the experimental design is a between-subjects one which permits investigations of subject-wise differentiations of retaliation intensities. Also, the way of presenting the rule of game is somehow different in the experiments; a (quite large) maximum number of rounds in each repeated game is announced to the subjects. The experimental results are mostly consistent with the results presented in the current paper; however, more detailed analysis are performed with the aid of the between-subjects design, thereby further implications are presented.

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

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

Ely, J., Välimäki, J., 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., Levine, D., Maskin, E., 1994. The Folk Theorem with Imperfect Public Information. Econometrica. 62, 997–1040.

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

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

Green, E., Porter, R., 1984. Noncooperative Collusion under Imperfect Price Information. Econometrica. 52, 87–100.

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

Kayaba, Y., Matsushima, H., Toyama, T., 2018. Accuracy and Retaliation in Repeated Games with Imperfect Private Monitoring: Experiments. Discussion Paper, CARF-F-433, University of Tokyo.

Mailath, J., Samuelson, L., 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., 2019. Behavioral Theory of Repeated Prisoner's Dilemma: Generous Tit-For-Tat Strategy. Discussion Paper, CARF-F-452, University of Tokyo.

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

Matsushima, H., Toyama, T., 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., Toyama, T., Yagi, N., 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., Sigmund, K., 1992. Tit-For-Tat in Heterogeneous Populations Nature. 355, 250–253.

Osborne, M., Rubinstein, A., 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.

Wu, J., Axelrod, R., 1995. How to Cope with Noise in the Iterated Prisoner's Dilemma. Journal of Conflict Resolution. 39, 183–189.

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 (sequence of game lengths)
October 2, 2018 (morning)	28	0.9 (11, 34, 48, 21, 33)
October 3, 2018 (morning)	28	0.9 (42, 11, 30, 46, 24)
October 3, 2018 (afternoon)	28	0.7 (11, 34, 48, 21, 33)
October 4, 2018 (morning)	28	0.7 (42, 11, 30, 46, 24)
June 11, 2019 (morning)	28	0.9 (7, 28, 45, 40, 28)
June 11, 2019 (afternoon)	28	0.9 (17, 44, 17, 31, 49)
June 12, 2019 (morning)	28	0.7 (7, 28, 45, 40, 28)
June 14, 2019 (morning)	28	0.7 (17, 44, 17, 31, 49)

Table 3:
Decisions and Signal

	$p = 0.9$			$p = 0.7$		
	N	Mean	St. Dev.	N	Mean	St. Dev.
Cooperative choice	16,968	0.845	0.362	16,968	0.598	0.490
Good signal	16,968	0.775	0.418	16,968	0.537	0.499

Table 4:
Means of Cooperative Action Choice

	$p = 0.9$	$p = 0.7$	p-values
$q(p)$ (round 1)	0.905 (0.019)	0.748 (0.033)	< 0.001
$r(c;p)$	0.924 (0.009)	0.726 (0.025)	< 0.001
Individual-level means	0.918 (0.010)	0.706 (0.026)	< 0.001
$r(d;p)$	0.560 (0.030)	0.437 (0.026)	0.002
Individual-level means	0.621 (0.024)	0.460 (0.026)	< 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 5:
Retaliation Intensities

	Mean	*S.E.*	*p-value*
$r(c;0.9)-r(d;0.9)$	0.365	0.027	< 0.001
Individual-level means	0.297	0.021	0.003
$r(c;0.7)-r(d;0.7)$	0.289	0.019	< 0.001
Individual-level means	0.247	0.020	< 0.001
$(r(c;0.9)-r(d;0.9))-(r(c;0.7)-r(d;0.7))$	0.075	0.033	< 0.001 (0.023+)
Individual-level means	0.050	0.029	< 0.001 (0.079+)

Notes: The standard errors are block-bootstrapped (subject and repeated game level) with 5,000 repetitions, which is used to calculate p-values for the hypothesis tests for the comparison to the value predicted by the theory ($w(p)$), which is 0.235 in the high accuracy treatment and 0.47 in the low accuracy treatment. Precisely, the null hypothesis is that the mean is identical to the value predicted by the theory.
+ The p-values shown in the parentheses are for hypothesis tests in which the null is that the mean is identical to zero.

Table 6:
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 7:
Maximum Likelihood Estimates of Individual Strategies

	p = 0.9	S.E.	p = 0.7	S.E.
ALL-C (g-TFT-1-1)	0.072**	(0.036)	0.118***	(0.035)
g-TFT-1-0.875	0	(0.027)	0.015	(0.017)
g-TFT-1-0.75	0.131***	(0.049)	0.037	(0.026)
g-TFT-1-0.625	0	(0.041)	0	(0)
g-TFT-1-0.5	0.040	(0.037)	0	(0)
g-TFT-1-0.375	0.023	(0.031)	0.006	(0.011)
g-TFT-1-0.25	0.021	(0.026)	0.009	(0.016)
g-TFT-1-0.125	0.032	(0.023)	0	(0.007)
TFT (g-TFT-1-0)	0	(0.008)	0.018	(0.013)
Random-0.875 (g-TFT-0.875-0.875)	0.015	(0.024)	0.003	(0.006)
g-TFT-0.875-0.75	0.098**	(0.042)	0.012	(0.024)
g-TFT-0.875-0.625	0	(0.034)	0.070*	(0.042)
g-TFT-0.875-0.5	0.134**	(0.054)	0.047	(0.042)
g-TFT-0.875-0.375	0.011	(0.023)	0.046	(0.042)
g-TFT-0.875-0.25	0	(0.008)	0.060	(0.037)
g-TFT-0.875-0.125	0.012	(0.011)	0	(0.009)
g-TFT-0.875-0	0	(0.006)	0	(0)
Random-0.75 (g-TFT-0.75-0.75)	0	(0)	0	(0)
g-TFT-0.75-0.625	0	(0)	0	(0.005)
g-TFT-0.75-0.5	0	(0.009)	0.072	(0.050)
g-TFT-0.75-0.375	0.018	(0.016)	0.087*	(0.050)
g-TFT-0.75-0.25	0	(0)	0	(0.011)
g-TFT-0.75-0.125	0	(0.006)	0	(0)
g-TFT-0.75-0	0	(0)	0	(0)
Random-0.625 (g-TFT-0.625-0.625)	0	(0)	0	(0)
g-TFT-0.625-0.5	0.006	(0.009)	0.019	(0.020)
g-TFT-0.625-0.375	0	(0)	0	(0.013)
g-TFT-0.625-0.25	0	(0)	0.044**	(0.023)
g-TFT-0.625-0.125	0	(0)	0	(0.004)
g-TFT-0.625-0	0	(0.012)	0	(0)
Random-0.5 (g-TFT-0.5-0.5)	0.012	(0)	0.004	(0.007)
g-TFT-0.5-0.375	0	(0)	0	(0)
g-TFT-0.5-0.25	0	(0)	0.004	(0.019)
g-TFT-0.5-0.125	0	(0)	0	(0.010)
g-TFT-0.5-0	0	(0)	0	(0)
Random-0.375 (g-TFT-0.375-0.375)	0	(0)	0	(0)
g-TFT-0.375-0.25	0	(0)	0	(0.011)
g-TFT-0.375-0.125	0	(0)	0.014	(0.017)
g-TFT-0.375-0	0	(0)	0	(0)
Random-0.25 (g-TFT-0.25-0.25)	0	(0)	0.028	(0.018)
g-TFT-0.25-0.125	0	(0)	0.024	(0.018)
g-TFT-0.25-0	0	(0)	0	(0)
Random-0.125 (g-TFT-0.125-0.125)	0	(0)	0	(0.006)
g-TFT-0.125-0	0	(0)	0	(0)
ALL-D (g-TFT-0-0)	0	(0)	0.045**	(0.020)
g-TF2T-0.875	0	(0.020)	0	(0.011)
g-TF2T-0.75	0	(0.027)	0.018	(0.023)
g-TF2T-0.625	0	(0.058)	0.003	(0.030)
g-TF2T-0.50	0.156**	(0.077)	0.044	(0.040)
g-TF2T-0.375	0	(0.040)	0.027	(0.042)
g-TF2T-0.25	0	(0.014)	0	(0.017)

g-TF2T-0.125	0	(0.005)	0	(0)
TF2T (g-TF2T-0)	0.050	(0.030)	0	(0.002)
TF3T	0.028	(0.031)	0	(0.009)
g-2TFT-0.875	0	(0)	0	(0)
g-2TFT-0.75	0	(0)	0	(0)
g-2TFT-0.625	0	(0)	0	(0)
g-2TFT-0.50	0	(0)	0	(0.002)
g-2TFT-0.375	0	(0)	0	(0)
g-2TFT-0.25	0	(0)	0	(0.007)
g-2TFT-0.125	0	(0)	0	(0.006)
2TFT (g-2TFT-0)	0	(0)	0	(0)
2TF2T	0.014	(0.016)	0.032	(0.030)
Grim	0	(0)	0.009	(0.007)
Grim-2	0.044*	(0.024)	0.011	(0.018)
Grim-3	0.084*	(0.046)	0.076**	(0.033)
Gamma	0.271***	(0.017)	0.351***	(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 8:
Frequency of Cooperative Strategies, g-TFT, Retaliation Intensities (RI)

	p = 0.9		p = 0.7	
Cooperative strategies		98.8%		88.2%
Non-cooperative strategies		1.2%		11.8%
Family of g-TFT		62.5%		78.1%
Excluding signal non-contingent, zero RI strategies (ALL-C, ALL-D, Random-*r*)		52.6%		58.5%
Equilibrium RI	(RI = 0.25)	13.1%	(RI = 0.5)	4.6%
Stronger RI	(RI > 0.25)	29.1%	(RI > 0.5)	9.3%
Weaker RI (positive RI only)	(0 < RI < 0.25)	10.4%	(0 < RI < 0.5)	44.6%
Including zero RI strategies	(RI < 0.25)	20.4%	(RI < 0.5)	64.2%
Mean RI conditional on g-TFT		0.312		0.274
Median RI conditional on g-TFT		0.250		0.250
Long-term punishment strategies		14.2%		12.8%
Long-memory (lenient) strategies		37.5%		21.0%

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 6) includes stochastic strategies, the choice probabilities should be extended to cover stochastic cases. Following Fudenberg, Rand, and Dreber (2012), (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. Note 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 Appendices C-G
Accuracy and Retaliation
in Repeated Games with Imperfect Private Monitoring:
Experiments

Online Appendix C: Screens in Constant Random Termination

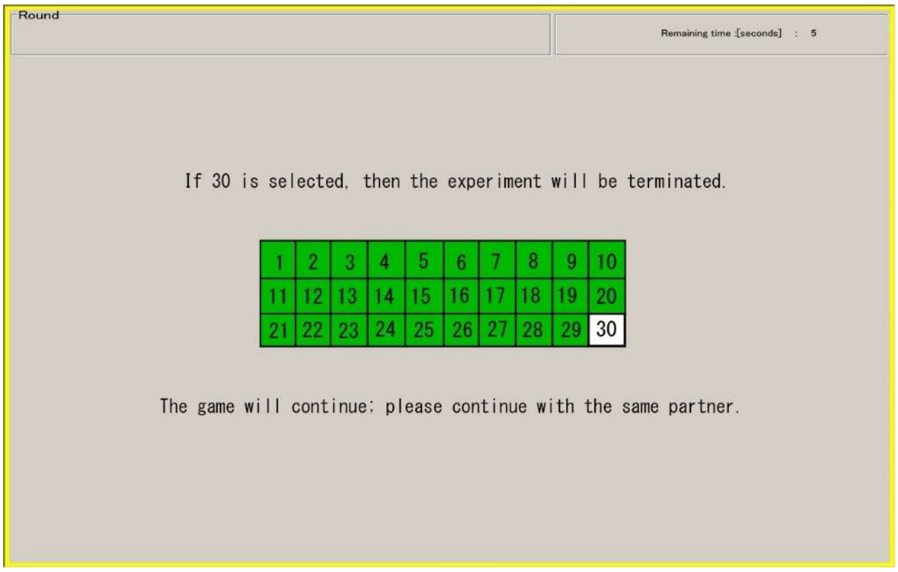


Figure C.1: Screen when the repeated game continues

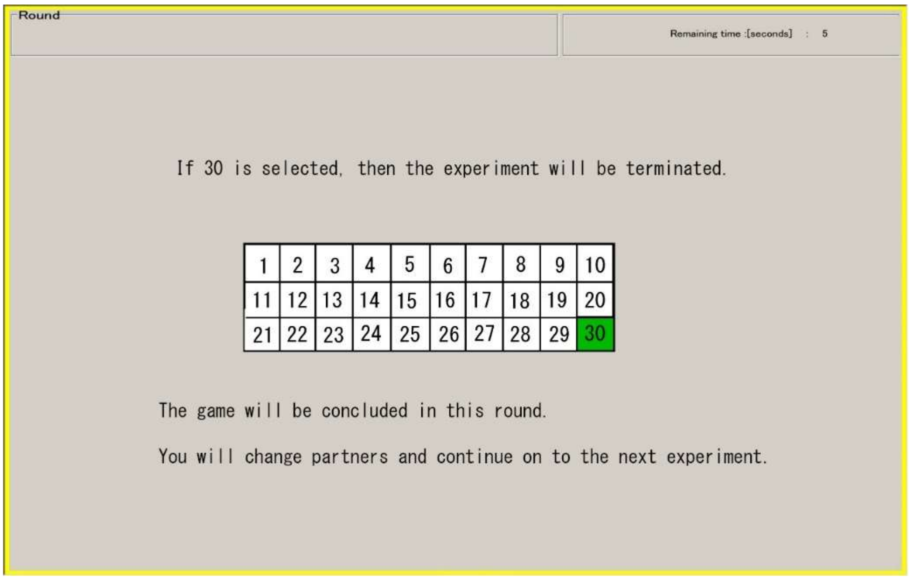


Figure C.2: Screen when the repeated game is terminated

Online Appendix D: Impact of Experience

Several studies have reported that the frequency of cooperation changes as people experience playing repeated games (see discussion in Dal Bó and Fréchette, 2016). Cooperation rates may rise with the experience of repeated games when the experimental parameters are conducive to cooperation. We document the learning effects observed in the data in detail in this appendix.

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 four explanatory variables, *RG2*, *RG3*, *RG4*, and *RG5*. The dummy variable *RG2* takes a value of 1 if the choice is made in the second repeated game in each treatment. Similarly, the dummy variables *RG3, RG4,* and *RG5* are defined with respect to the third, fourth, and fifth repeated games. The dummy variable for the first repeated game is omitted for non-singularity. 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. The regression coefficients on the dummy variables for repeated games capture the additional impacts on cooperation rates in comparison to that in the first repeated game.

[TABLE D.1 HERE]

Table D.1 displays the regression results for over-all cooperation rates. In the high accuracy treatment, the coefficients of the dummy variables for repeated games in the early stage increase as the experiments proceed. The growing positive values of the coefficients indicate that the subjects tend to cooperate more with experience, indicating that the welfare of the two players improves with experience. However, the cooperation rates stop growing in the final three repeated games; the coefficient of the dummy variable *RG3,* which is 0.119, is not statistically significantly different from both of the two coefficients of the dummy variables *RG4* and *RG5*, which are 0.139 (F-test, $p = 0.302$) and 0.131 ($p = 0.565$), respectively.

A similar pattern of learning is observed in the signal contingent cooperation rates. Table D.2 displays the regression results for the signal contingent cooperation rates conditional on good signal, and Table D.3. displays them conditional on bad signal. The coefficients for dummy variables for repeated games in the early stage become larger as the subjects experience games in the high accuracy treatment; however, the enlargement stops in the final three repeated games. For the signal contingent cooperation rates on good signal, the cooperation rates gain about 7% until the third repeated games; however, no additional gain is observed in the final three repeated games; the coefficient of the dummy variable *RG3,* which is 0.070, does not statistically significantly differ from both of the two coefficients of *RG4* and *RG5*, which are 0.076 ($p = 0.586$) and 0.079 ($p = 0.913$), respectively. For the bad signal, the cooperation rates gain about 15% until the third repeated games; however, the gain has insignificant difference in the final three repeated games; the coefficient of *RG3,* which is 0.149, does not statistically significantly differ from both of the two coefficients of *RG4* and *RG5*, which are 0.180 ($p = 0.473$) and 0.181 ($p = 0.450$), respectively.

[TABLE D.2 HERE]

[TABLE D.3 HERE]

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 of the dummy variable captures the contrast between the cooperation rate contingent on the good and bad signals, that is, the retaliation intensity. To examine learning effects on retaliation intensities across repeated games, we add the cross-product terms of *Signal* and the dummy variables for the repeated games, *RG2*, *RG3, RG4*, and *RG5,* in the sets of explanatory variables. The coefficients of the cross products terms capture the additional impacts on retaliation intensity in comparison to that in the first repeated game. 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 D.4 HERE]

Table D.4 displays the regression results. The regression coefficients in the high accuracy treatment indicate possible negative learning effect in the final three repeated games. The point estimates for the coefficients on the cross-product terms of *Signal* and *RG3, RG4, and RG5* are negative (though that for *RG3* is insignificant, and that for *RG5* is only marginally significant). However, these values do not statistically significantly differ among them; the coefficient on the cross-product terms of *Signal* and *RG3,* which is -0.044, does not statistically significantly differ from both of the coefficients on the two cross-product terms with respect to *RG4* and *RG5*, which are -0.103 ($p = 0.177$) and -0.075 ($p = 0.517$) respectively.

The discussion about the learning effects in the high accuracy treatment demonstrated above indicates that, we observe some learning effects in the early stage of the experiments, however additional learning is almost none or insignificantly small in the final three repeated games.

The stability of the cooperation rates in the final three repeated games are also observed in the low accuracy treatment. Rather, even in the early stage, learning effects are generally smaller in size or often not observed in the low accuracy treatment. As displayed in Table D.1, the absolute sizes of the regression coefficients in the low accuracy treatment tend to be small. The only significant coefficient is the one on the dummy variable *RG5*, which is 0.077. Again, additional learning is almost none or insignificantly small in the final three repeated games; the coefficient on the dummy variable *RG3,* which is 0.041, does not statistically significantly differ from both coefficients of the dummy variables, *RG4* and *RG5*, which are 0.034 ($p = 0.822$) and 0.077 ($p = 0.157$) respectively. Qualitatively similar results are observed in the signal contingent cooperation rates displayed in Tables D.2 and D.3. Moreover, for the retaliation intensities, none of the coefficients on the cross-product terms of *Signal* and *RG3, RG4, and RG5* differ from zero statistically. This indicates that there are no remarkable learnings for the retaliation intensities in the low accuracy treatment.

Studies on repeated games frequently report that cooperation often increases with experience when the experimental parameters are favorable for cooperation. Our results that positive learning in cooperation rates is observed in the early stages of the experiments in the high accuracy treatment are consistent with existing findings in the literature, although the size of the learning effects is not as large as observed in them.

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

	p = 0.9	**p = 0.7**
RG2	0.080**(0.032)	-0.051 (0.031)
RG3	0.119*** (0.032)	0.041 (0.032)
RG4	0.139*** (0.026)	0.034 (0.030)
RG5	0.131*** (0.027)	0.077**(0.032)
Observations	16,968	16,968
R2	0.015	0.008

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. The coefficient of *RG3* does not differ significantly from that of *RG4* in both the treatments (F-test, $p = 0.302$ for the high accuracy treatment and $p = 0.822$ in the low accuracy treatment); the coefficient of *RG3* does not differ significantly from that of *RG5* in both treatments ($p = 0.565$ for the high accuracy treatment and $p = 0.157$ in the low accuracy treatment). Furthermore, the coefficient on *RG4* does not differ significantly from that of *RG5* in the high accuracy treatment ($p = 0.643$), and it differs only marginally in the low accuracy treatment ($p = 0.092$).

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

	p = 0.9	p = 0.7
RG2	0.051*** (0.018)	-0.054* (0.032)
RG3	0.070*** (0.018)	0.043 (0.030)
RG4	0.076*** (0.017)	0.048 (0.030)
RG5	0.069*** (0.018)	0.065**(0.031)
Observations	12,721	8,829
R2	0.008	0.009

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. The coefficient of *RG3* does not differ significantly from that on *RG4* in both treatments (F-test, $p = 0.586$ for the high accuracy treatment and $p = 0.839$ in the low accuracy treatment); the coefficient of *RG3* does not differ significantly from that of *RG5* in both treatments ($p = 0.913$ for the high accuracy treatment and $p = 0.326$ in the low accuracy treatment). Furthermore, the coefficient of *RG4* does not differ significantly from that of *RG5* in both treatments ($p = 0.480$ for the high accuracy treatment and $p = 0.493$ in the low accuracy treatment).

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

	p = 0.9	**p = 0.7**
RG2	0.115** (0.046)	-0.049 (0.032)
RG3	0.149*** (0.052)	0.039 (0.034)
RG4	0.180*** (0.043)	0.032 (0.032)
RG5	0.181*** (0.044)	0.071**(0.035)
Observations	3,678	7,579
R2	0.014	0.006

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. The coefficient of *RG3* does not differ significantly from that on *RG4* in both treatments (F-test, $p = 0.450$ for the high accuracy treatment and $p = 0.812$ in the low accuracy treatment); the coefficient of *RG3* does not differ significantly from that of *RG5* in both treatments ($p = 0.473$ for the high accuracy treatment and $p = 0.260$ in the low accuracy treatment). Furthermore, the coefficient of *RG4* does not differ significantly from that of *RG5* in both treatments ($p = 0.998$ for the high accuracy treatment and $p = 0.147$ in the low accuracy treatment).

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

	p = 0.9	p = 0.7
Signal	0.360*** (0.031)	0.257*** (0.035)
Signal: RG2	0.007 (0.041)	-0.020 (0.035)
Signal: RG3	-0.044 (0.050)	-0.017 (0.030)
Signal: RG4	-0.103** (0.040)	0.007 (0.032)
Signal: RG5	-0.075* (0.038)	-0.018 (0.033)
RG2	0.054 (0.047)	-0.040 (0.033)
RG3	0.118** (0.056)	0.048 (0.033)
RG4	0.184*** (0.042)	0.029 (0.031)
RG5	0.157*** (0.041)	0.078** (0.036)
Observations	16,408	16,408
R2	0.186	0.093

Notes: Cluster-robust (individual-level) standard errors in parenthesis. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. The coefficient of the cross-product term *Signal*: *RG3* does not differ significantly from that of *Signal*: *RG4* in both treatments (F-test, $p = 0.177$ for the high accuracy treatment and $p = 0.352$ in the low accuracy treatment); the coefficient on the cross-product term *Signal*: *RG3* does not differ significantly from that of *Signal*: *RG5* in both treatments ($p = 0.517$ for the high accuracy treatment and $p = 0.946$ in the low accuracy treatment). Furthermore, the coefficient on the cross-product term *Signal*: *RG4* does not differ significantly from that of *Signal*: *RG5* in both treatments ($p = 0.381$ for the high accuracy treatment and $p = 0.330$ in the low accuracy treatment).

Online Appendix E: The Validity of SFEM with a Large Set of g-TFT Strategies

Our study was the first to include a large set of mixed strategies in the strategy list. The simulation exercises provided by Fudenberg, Rand, and Dreber (2012) included mainly deterministic strategies with a few g-TFT strategies in the strategy list in order to investigate whether the SFEM could distinguish memory-one strategies from long-memory (lenient) strategies. Similarly, we performed simulation exercises in order to investigate the efficacy of our SFEM, where the strategy list included an extensive set of g-TFT strategies in our experimental environment.

In accordance with the main claims of the study, the aims of the simulation exercises here were as follows:

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

In the following simulations, mimicking our experimental environments, we generated 112 subjects for each treatment; each subject experienced three interactions, which mimicked the final three repeated games used in the SFEM. Each interaction included 30 rounds. The partners were randomly altered across the interactions. The SFEM strategy list employed in the following simulations is the same as the one employed in the main study (Table 5). We repeated the data generation and estimation processes 100 times for each simulation, and we reported the means and standard deviations of the estimates.

A. High accuracy case

Here, we set up the following three types of players:

a. G-TFT-0.875-0.5, which had the largest individual share in our SFEM estimates in the high accuracy treatment

b. TFT, which examined whether the SFEM correctly distinguished g-TFT-0.875-0.5 from TFT
c. TF2T, which examined whether the SFEM correctly distinguished the long-memory strategy from the strategies in the g-TFT class strategies (i.e., memory-one strategies)

In this simulation, approximately one-third of the total subjects (37 subjects) played g-TFT-0.875-0.5, another one-third (37 subjects) played TFT, and the rest (38 subjects) played TF2T. The signal accuracy was set to match the higher accuracy of the study (i.e., $p = 0.9$).

[TABLE E.1 HERE]

Table E.1 summarizes the results of the SFEM. Each estimate for TFT and TF2T was found to be accurate. The estimate for TFT was 0.3304, which was identical to the true value (0.3304 or 37/112). Similarly, the estimate for TF2T was 0.3393, which, again, was identical to the true value (0.3393 or 38/112). The SFEM accurately distinguished the deterministic strategies from the mixed strategies and distinguished the long-memory strategies from the strategies in the class of g-TFT strategies (i.e., memory-one strategies).

The estimate for g-TFT-0.875-0.5 had some biases. The estimated value was 0.2957, which was smaller than the true value by approximately 0.035 (true: 0.3304). Instead, similar "nearby" g-TFT strategies (g-TFT-0.85-0.375, g-TFT-0.875-0.625, …etc.) had some small values in the estimates (true: 0 for each). However, the maximum value among them was, at the most, 0.0152; furthermore, none of them statistically significantly differed from zero.

Since the bias occurred only among "nearby" g-TFT strategies, its effect on the estimate regarding the total fraction of g-TFT strategies and on the estimate regarding the mean of retaliation intensities should have been, at the most, marginal. Indeed, the total fraction of the strategies in a class of g-TFT strategies was estimated accurately: 0.6607 (true: 0.6607 or 74 out of 112, consisting of g-TFT-0.875-0.5 and TFT). Furthermore, the mean of retaliation intensities among g-TFT players was estimated approximately. The estimated value was 0. 6685; this was slightly larger than the true value (0.6670) by only 0.0015.

B. Low accuracy case

Here, we set up the following three types of players:

a. G-TFT-0.75-0.375, which had the largest individual share in our SFEM estimates in the low accuracy treatment
b. TFT, which examined whether the SFEM correctly distinguished g-TFT-0.75-0.375 from TFT
c. TF2T, which examined whether the SFEM correctly distinguished the long-memory strategies from the g-TFT class strategies (i.e., memory-one strategies)

In this simulation, approximately one-third of the total subjects (37 subjects) played g-TFT-0.75-0.375, another one-third (37 subjects) played TFT, and the rest (38 subjects) played TF2T. The signal accuracy was set to the lower accuracy of the study (i.e., $p = 0.7$).

[TABLE E.2 HERE]

Similar successful results were obtained in the low accuracy treatment. Table E.2 summarizes the results of the SFEM. Again, each estimate for TFT and TF2T was found to be accurate. The estimate for TFT was 0.3304, which was identical to the true value (0.3304 or 37/112). Similarly, the estimate for TF2T was 0.3393, which, again, was identical to the true value (0.3393 or 38/112). The SFEM accurately distinguished the deterministic strategies from the mixed strategies and distinguished the long-memory strategies from the strategies in the class of g-TFT strategies (i.e., memory-one strategies).

The estimate for g-TFT-0.75-0.375 had some biases. The estimate was 0.3032, which is smaller than the true value by approximately 0.03 (true: 0.3304). Instead, similar “nearby” g-TFT strategies (g-TFT-0.75-0.5, g-TFT-0.75-0.25, g-TFT-0.875-0.375,

…etc.) had some small positive values in the estimates (true: 0 for each). However, the maximum value among them was, at the most, 0.0071; none of them statistically significantly differed from zero.

Since the bias occurred only among "nearby" g-TFT strategies, its effect on the estimate for the total fraction of g-TFT strategies and on the estimate for the mean of the retaliation intensities should have been, at the most, marginal. Indeed, the total fraction of the strategies in the class of g-TFT strategies was estimated correctly: 0.6607 (true: 0.6607 or 74/112, consisting of g-TFT-0. 0.75-0.375 and TFT). Furthermore, the mean of retaliation intensities among g-TFT players was estimated approximately; the estimate was 0. 0.6686, which was slightly larger than the true value (0.6670) by only 0.0016.

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

	Estimate	True Value
Generous-TFT-0.875-0.75	0.0018 (0.0064)	0
Generous-TFT-0.875-0.625	0.0150 (0.0258)	0
Generous-TFT-0.875-0.5	0.2957 (0.0335)	0.3304 (=37/112)
Generous-TFT-0.875-0.375	0.0152 (0.0242)	0
TFT	0.3304 (0.0000)	0.3304 (=37/112)
TF2T	0.3393 (0.0000)	0.3393 (=38/112)
Other strategies (each)	Less than 0.001	0

Notes: Standard deviations are provided in parenthesis.

Table E.2:
The Mean of SFEM Estimates for the Simulated Data in the High Accuracy Case (p = 0.7)

	Estimate	True Value
Generous-TFT-0.875-0.500	0.0011 (0.0046)	0
Generous-TFT-0.875-0.375	0.0066 (0.0116)	0
Generous-TFT-0.875-0.250	0.0016 (0.0042)	0
Generous-TFT-0.750-0.500	0.0070 (0.0132)	0
Generous-TFT-0.750-0.375	0.3032 (0.0246)	0.3304 (=37/112)
Generous-TFT-0.750-0.250	0.0071 (0.0141)	0
Generous-TFT-0.625-0.250	0.0011 (0.0035)	0
TFT	0.3304 (0.0000)	0.3304 (=37/112)
TF2T	0.3393 (0.0000)	0.3393 (=38/112)
Other strategies (each)	Less than 0.001	0

Notes: Standard deviations are provided in parenthesis.

Online Appendix F: Robustness Checks for Our Strategy Estimation

In this part of the appendix, we will discuss the robustness checks for the SFEM estimates. In the main text, we used all three repeated games of each treatment in our estimation. Here, we demonstrated that the estimation results showed almost no changes, even when the final four repeated games were utilized in each treatment (instead of the three utilized for the main results) (Tables F.1 and F.2).

[TABLE F.1 HERE]

[TABLE F.2 HERE]

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

	Final three	Final four
Cooperative strategies	98.8%	100.0%
Family of g-TFT (including zero RI strategies)	62.5%	65.7%
Equilibrium RI	13.1%	13.6%
Stronger RI	29.1%	30.9%
Weaker RI (including zero RI strategies)	20.4%	21.2%
Mean RI conditional on g-TFT (including zero RI strategies)	0.312	0.320
Long-term punishment strategies	14.2%	8.9%
Long-memory (lenient) strategies	37.5%	34.3%

Table F.2:
Aggregated Estimates for the Low Accuracy Case (p = 0.7)

	Final three	Final four
Cooperative strategies	88.2%	86.7%
Family of g-TFT (including zero RI strategies)	78.1%	71.1%
Equilibrium RI	4.6%	1.2%
Stronger RI	9.3%	10.0%
Weaker RI (including zero RI strategies)	64.2%	59.8%
Mean RI conditional on g-TFT (including zero RI strategies)	0.274	0.285
Long-term punishment strategies	12.8%	15.7%
Long-memory (lenient) strategies	21.0%	28.0%

Online Appendix G:
Experimental Instruction and Computer Screen Images
(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. Instructions – 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.

Your partner / You	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.

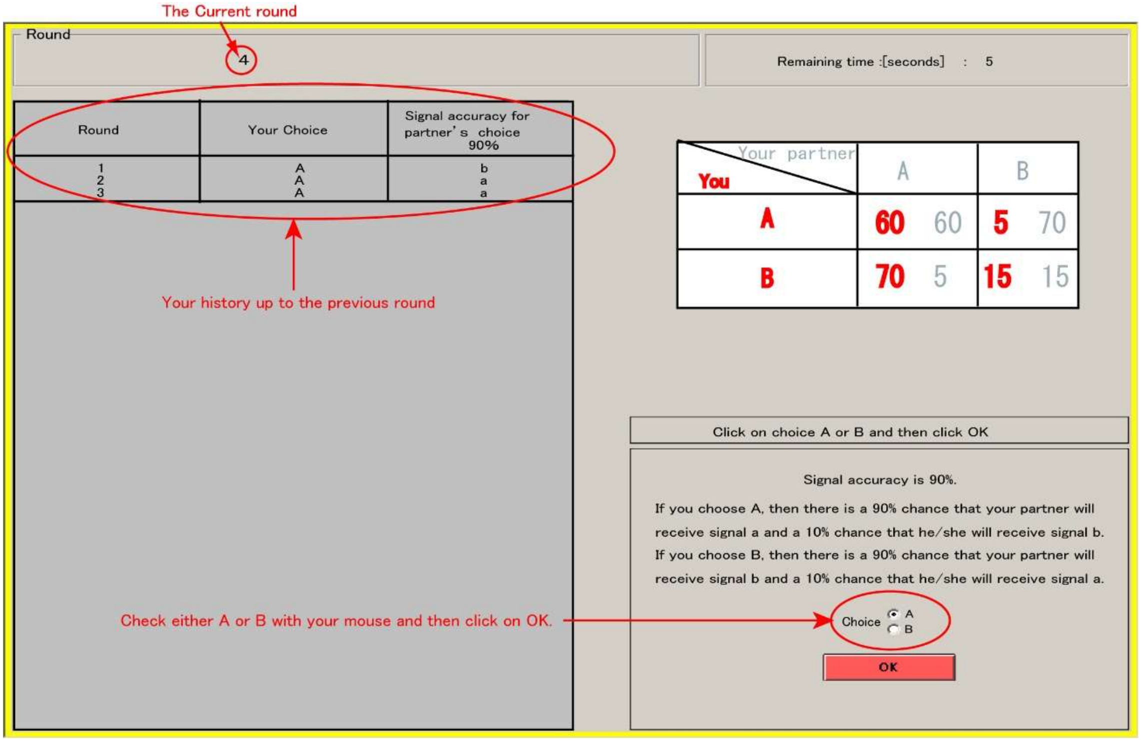
Please make sure that you fill out the questionnaire and the bank transfer form correctly.
Please raise your hand quietly if you have any questions.

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.

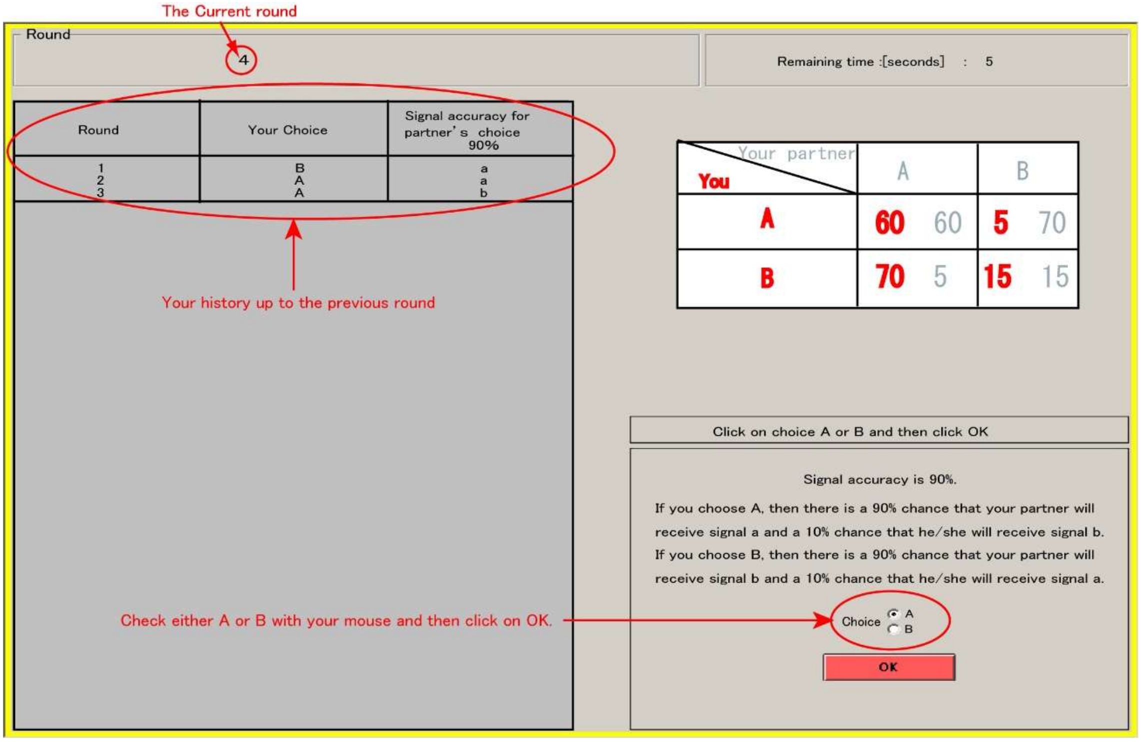
Please do not disclose any details regarding the experiments to anyone. Thank you very much for your participation. Please follow the instructions of the experimenters to leave the room.

2. Computer Screen Images

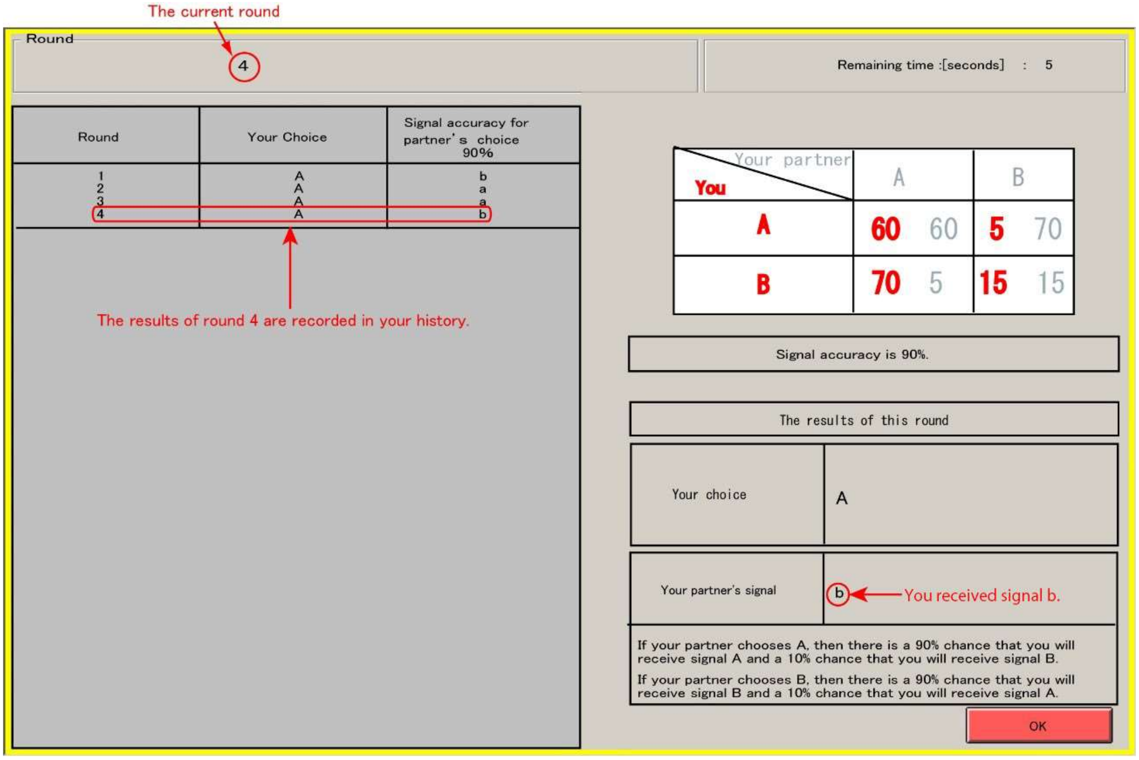
Screen 1: Your Selection Screen



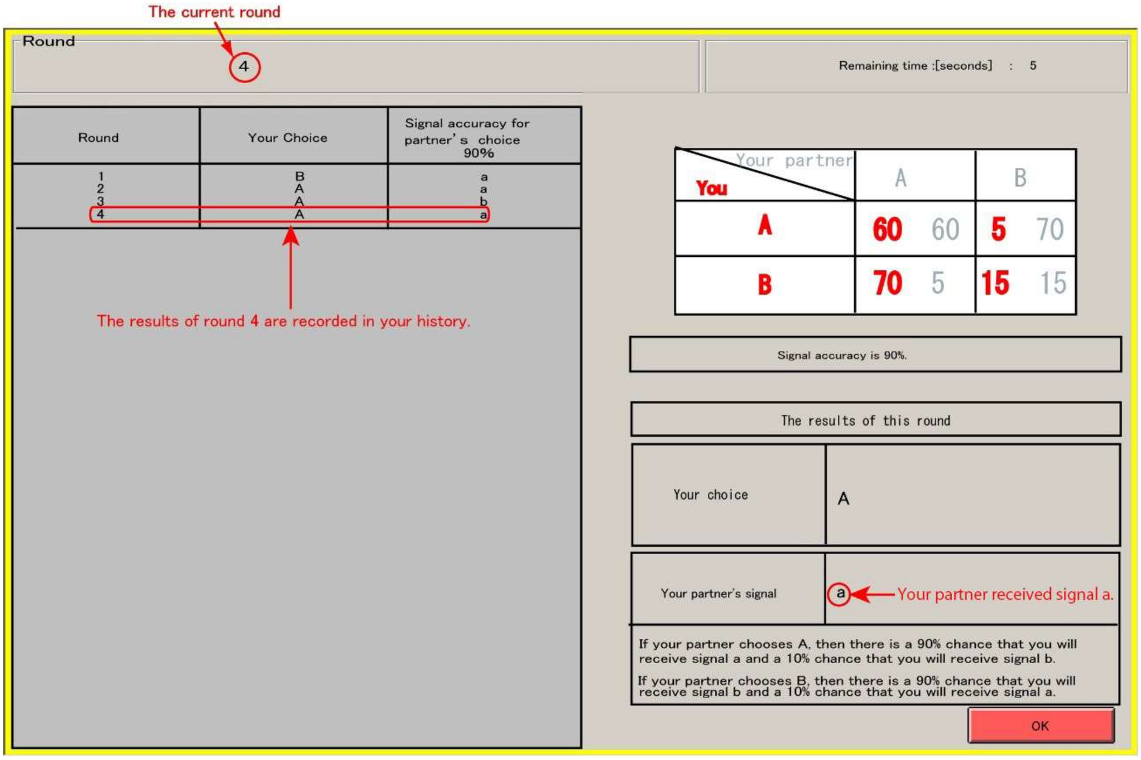
Screen 2: Your Partner's Selection Screen



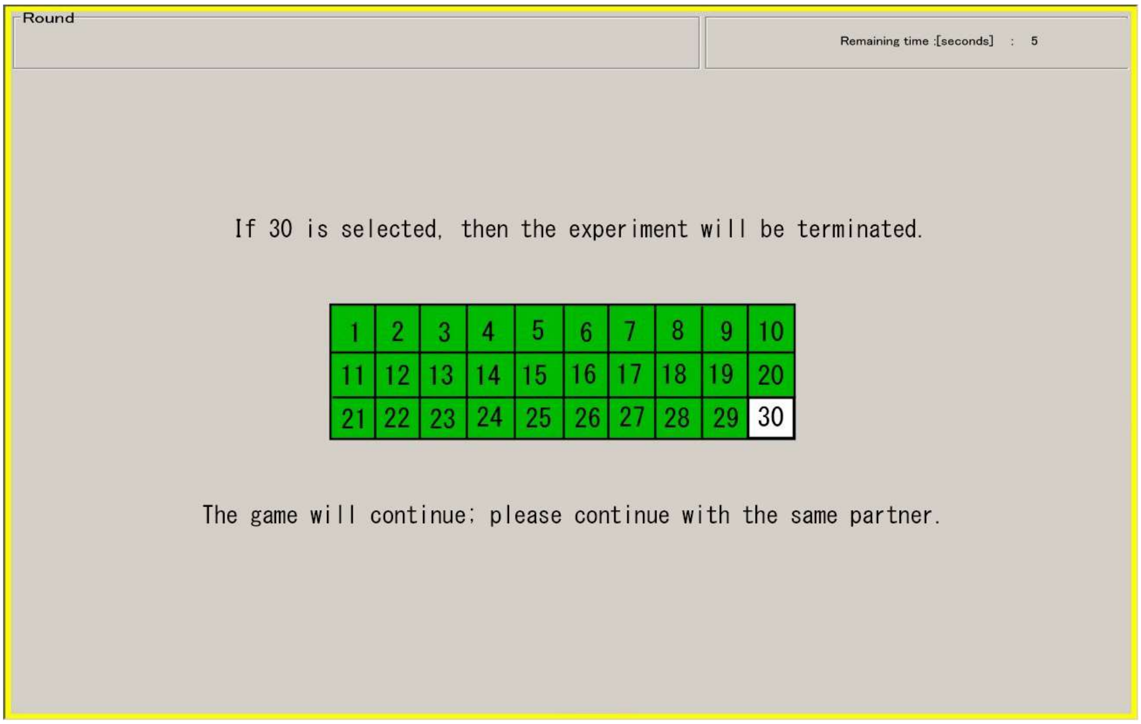
Screen 3: Your Results Screen



Screen 4: The Results Screen of Your Partner



Screen 5: Lottery (experiment continues)



Screen 6: Lottery (experiment is terminated)