CIRJE-F-879

Behavioral Approach to Repeated Games with Private Monitoring

Hitoshi Matsushima
University of Tokyo

Tomomi Tanaka
Economic Development & Global Education, LLC

Tomohisa Toyama
Faculty of Engineering, Kogakuin University

March 2013

Behavioral Approach to Repeated Games with Private Monitoring[1]

Hitoshi Matsushima
Department of Economics, University of Tokyo[2]

Tomomi Tanaka
Economic Development & Global Education, LLC

Tomohisa Toyama
Faculty of Engineering, Kogakuin University

First Version: April 2, 2011
This Version: March 12, 2013

Abstract

We examine repeated prisoners' dilemma with imperfect private monitoring and random termination where the termination probability is low. We run laboratory experiments and show subjects retaliate more severely when monitoring is more accurate. This experimental result contradicts the prediction of standard game theory. Instead of assuming full rationality and pure self-interest, we introduce naiveté and social preferences, i.e., reciprocal concerns, and develop a model that is consistent with, and uniquely predicts, the observed behavior in the experiments. Our behavioral model suggests there is a trade-off between naiveté and reciprocity. When people are concerned about reciprocity, they tend to make fewer random choices.

JEL Classification Numbers: C70, C71, C72, C73, D03.
Keywords: Infinitely Repeated Prisoners' Dilemma, Imperfect Private Monitoring, Experimental Economics, Monitoring Accuracy, Social Preference, Generous Tit-for-Tat Strategy.

---

[1]This paper is a revised version of Matsushima and Toyama (2011). Substantial contents of this paper concerning experimental analyses and behavioral models are not part of that work. We are grateful to Mr. Nobuyuki Yagi for helping us run the experiments. Research for this paper was supported by a Grant-in-Aid for Scientific Research (Kakenhi 21330043) 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 and a grant from the 21st Century Culture, Arts, and Sciences Foundation.
[2] Department of Economics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: hitoshi at mark e.u-tokyo.ac.jp.

1. Introduction

It's a well-accepted view in the literature on theory of repeated games with perfect monitoring that when a player deviates from the collusive relationship by selecting a defective action, his partner will retaliate by selecting a defective action. Hence, implicit collusion can be sustained as a subgame perfect equilibrium if the deviant's instantaneous gain is exceeded by a future loss caused by his partner's retaliation.[3]

The device of contingent action switch in the above manner can be applied to the case of imperfect monitoring, but with a limit.[4] Since a player cannot directly observe his partner's action choice, he instead makes his choice of action contingent on the observed signal. Since monitoring is imperfect, it's inevitable a player may observe a bad signal even when his partner has selected a cooperative action. This causes a welfare loss peculiar to imperfect monitoring; a player might retaliate even when his partner has actually selected a cooperative action.

Standard game theory with rational and purely self-interested players suggests improvements in monitoring accuracy can decrease a welfare loss without contradicting incentive constraints. If monitoring technology becomes more accurate, a player can induce their partner to choose a cooperative action by using a milder punishment i.e., by being less responsive to whether the signal is good or bad. To clarify this point, this paper compares the behavior observed in laboratory experiments with theoretical predictions induced by a simple form of Nash equilibrium, namely, symmetric generous tit-for-tat Nash equilibrium.[5] For simplicity, we assume a player only responds to the signal he has received in the previous round. Symmetric generous tit-for-tat Nash equilibrium predicts the less accurate the monitoring technology, the more severely a purely self-interested player retaliates against his partner.

---

[3] For the theoretical studies and surveys on infinitely repeated games with perfect monitoring, see Abreu (1988), Fudenberg and Tirole (1993), Osborne and Rubinstein (1994), and Mailath and Samuelson (2006). This view is supported by experimental studies such as Dal Bó, (2005).

[4] For example, see Green and Porter (1994) and Abreu, Pearce, and Stacchetti (1990).

[5] The concept of the generous tit-for-tat Nash equilibrium was explored by Nowak and Sigmund (1992), Takahashi (1997), Matsushima (2010), and others.

Our experimental results, however, do not support this prediction. When monitoring technology is more accurate, laboratory subjects tend to retaliate much more severely than this equilibrium predicts.

We regard the difference in frequencies of cooperative action choices between the good signal and the bad signal as a proxy for the intensity of retaliation. Our experimental results suggest this difference tends to increase in accuracy, i.e., the subject react excessively to the signals as monitoring accuracy improves. This tendency makes the achievement of implicit collusion less likely to happen.

In order to explain the behavior observed in our experiments, we provide a new behavioral model that incorporates bounded rationality. We assume with some probability a player is naïve, i.e. he makes random choices between the cooperative and defective actions. With the remaining probability, he is sophisticated and is motivated not only by pure self-interests but also by social preferences, namely reciprocal concerns. He bears a psychological cost of guilt if he chooses a defective action after having received a good signal. On the other hand, he bears a psychological cost of resentment if he chooses a cooperative action after having received bad signals.[6]

We investigate an arbitrary accuracy-contingent generous tit-for-tat strategy profile that is consistent with the following observations from our experimental results; 1) subjects are more likely to select the cooperative action when they have observed good signals than bad signals. 2) Subjects select the cooperative action more often when the signal is more accurate. 3) The more accurate the signal technology, the greater the difference in frequency of cooperative action choices between good and bad signals. We permit a behavioral model to be accuracy-contingent in that the degrees of naiveté and reciprocity concerns are dependent on the accuracy levels.

We then show for any such accuracy-contingent generous tit-for-tat strategy profile, there is a unique accuracy-contingent behavioral model in which this profile is a Nash

[6] Economists often call the tendency to return kindness to whom they are treated kindly, and the tendency to punish people who treated them badly "positive reciprocity", and "negative reciprocity", respectively (Fehr and Gächter, 2000). We explicitly incorporate psychological costs associated with these reciprocal concerns.

equilibrium with modification, namely a quasi-Nash equilibrium, at all times. This theoretical finding characterizes the behavioral mode of a bounded-rational player in a manner that; 1) the probability of naiveté is single-peaked across monitoring accuracies, 2) There is a trade-off between naiveté and social preferences in that when people have reciprocity concerns, they tend to make fewer random choices. Similarly, when people are making random choices, they tend not to have strong reciprocity concerns. 3) When monitoring technology is more accurate, a player suffers from the psychological costs of resentment if he selects a cooperative action after having received a bad signal. This makes it hard for players to sustain the cooperative outcome in the case of high monitoring accuracy.

Our behavioral approach has strong predictive power not just in clarifying motives as a behavioral model but also in describing strategy as unique equilibrium. We use experimental data to identify behavioral patterns, and construct a behavioral model. By using experimental data, we restrict the candidates of equilibria out of numerous equilibria that exist in repeated games with imperfect private monitoring. We find there exists a unique quasi-Nash equilibrium that is consistent with the behavioral patterns.

The contribution of this paper is to demonstrate how the standard theory of infinitely repeated games with imperfect private monitoring does not explain behavior very well, and propose a behavioral model that is capable of explaining behavioral deviations from standard game theoretic predictions. Our model is behavioral in two ways. 1) We incorporate motives other than self-interests, i.e. naïve and reciprocity concerns. 2) Our model is context dependent in that the degrees of naiveté and reciprocal concerns are dependent on the accuracy level. We take into account the fact that people's motives and perceptions are often context dependent. Let us consider the situation in which individual's effort cannot be perfectly monitored by his partner. When monitoring is highly accurate, it may be harder to rebuild trust if your partner has received a bad signal about your effort level. However, if monitoring is poorly accurate, your partner may not punish you even if he has received a bad signal about your effort. Our behavioral model is capable of explaining such phenomenon.

This paper also makes an important contribution to the field of repeated games with imperfect private monitoring by offering a solution to a fundamental problem in this field. Repeated games with imperfect private monitoring have numerous equilibria, and identifying strategies and equilibria is considered a difficult task. However, by using experimental data, we are able to refine a unique quasi-Nash equilibrium that is consistent with behavioral patterns, and predict equilibrium strategies people may take.

The organization of this paper is as follows. In Section 2, we summarize related literature, and highlight the contribution of our study in the literature. In Section 3, we develop a model. Section 4 introduces the concept of symmetric generous tit-for-tat Nash equilibrium, and establishes the theorem that characterizes the class of symmetric generous tit-for-tat Nash equilibria. Section 5 illustrates the experimental design. Section 6 shows our experimental results. In Section 7, we incorporate bounded rationality in our model, propose a behavioral model, and refine a unique quasi-Nash equilibrium. Section 8 concludes.

2. Related Literatures

There is a large and still growing literature of repeated game theory with private monitoring that provides various folk-theorem-type results. This literature generally assumes players are rational and purely self-interested, and it is possible for a wide range of payoff vectors to be sustained by Nash equilibria even if monitoring is rather inaccurate, provided the discount factor is close to unity.[7] This literature, however, does not make any prediction on the equilibrium strategies people may actually take. In order to make the achievement of payoff vectors consistent with Nash equilibrium play, the authors generally

[7] For early studies on repeated games with private monitoring, see Radner (1986) and Matsushima (1990a, 1990b). The subsequent works such as Sekiguchi (1997), Ely and Välimäki (2002), and Piccione (2002) demonstrate the folk-theorem-like results when monitoring is private but almost perfect. Matsushima (2004) proves the folk theorem in the repeated prisoners' dilemma with conditional independence when monitoring is inaccurate. For surveys on the progress of repeated game theory with private monitoring, see Kandori (2002) and Mailath and Samuelson (2006). Sugaya (2012) recently demonstrates the general folk theorem.

tailor equilibrium strategies to the fine details of model specifications in a mathematical manner. This paper offers an alternative approach; we focus on a simple class of strategies, namely generous tit-for-tat strategies, and then explore the behavioral patterns observed in laboratory experiments, and offer a behavioral model consistent with the behavioral patterns by incorporating bounded rationality into the model. We then show there exists a unique quasi-Nash equilibrium consistent with the behavioral patterns.

Except for Aoyagi and Fréchette (2009), there are few studies that test infinitely repeated games with imperfect monitoring by laboratory experiments. Aoyagi and Fréchette (2009) run laboratory experiments to study infinitely repeated prisoners' dilemma with imperfect monitoring and vary the noise in monitoring. Their findings show laboratory subjects are sophisticated enough to increase the level of cooperation as monitoring accuracy improves. Aoyagi and Fréchette assume monitoring is public and the publicly observable signal is a single-dimensional real variable; from this signal, it is impossible to identify which player is more likely to deviate. Our substantial departure from Aoyagi and Fréchette is to shed light on the rise of bounded rationality if monitoring is private. Repeated games with imperfect private monitoring have many applications such as work efforts, intra-household resource allocations, and coordination of global environmental protection efforts. When monitoring is private, players' action choices are monitored through the observation of their respective private signals.[8] Since signals are private, people do not have common knowledge about the past actions, making it more likely that people choose random actions. For this reason, we take into account naiveté in our behavioral model.

Experimental research on various multi-stage models such as the ultimatum game, trust game, and gift exchange show social preference-based motives often facilitate cooperation in a one-shot game framework.[9] Subjects are often motivated by social preferences

[8] There are experimental studies, such as Holcomb and Nelson (1997) and Feinberg and Snyder (2002), in which subjects do not necessarily observe the same signals.
[9] This paper is also related to the literature of psychological games such as Geanakoplos, Pearce, and Stacchetti (1989), Rabin (1993), Charness and Dufwenberg (2005). This paper implicitly assumes that the dependence of preference on context such as the level of monitoring accuracy stem from each player's belief on how informative the observed signals are about the other players' past action choices

concerning reciprocity in ways that a player's friendly or hostile actions induce their partners to behave altruistically or unkindly, respectively.[10] Based on these experimental findings, it's possible to anticipate how social preferences could also facilitate cooperation in the repeated game framework. For these reasons, we take into consideration two types of bounded rationality in our behavioral model, namely naiveté and reciprocity.

Duffy and Muñoz-García (2012) demonstrate social preferences help the achievement of implicit collusions in infinitely repeated games and operate as substitutes for time discounting.[11] Duffy and Muñoz-García assume the discount factor is close to zero and monitoring is perfect. In contract, this paper assumes the discount factor is close to unity and monitoring is imperfect.

Several authors in the literature of social preferences assume preferences crucially depend on the context of the game being played.[12] This paper demonstrates a context-dependent behavioral model that describes how varying context influences the types and degrees of naiveté and reciprocal motives in a systematic and intuitive manner. In this paper, the relevant context to preferences is parameterized by the level of monitoring accuracy. We show the degree of monitoring accuracy influences the types and the degrees of reciprocal motives, and our behavioral model suggests there is a trade-off between naiveté and reciprocal concerns.

In sum, the contribution of this paper is the introduction of behavioral approach to repeated games with private monitoring. We offer an alternative way to refine equilibria, and provide predictive powers to repeated games with private monitoring. We also contribute to the social preference literature by extending its idea to the area of repeated games with private monitoring.

---

and also from his higher-order beliefs.

[10] For examples, see Güth, Schmittberger, and Schwarze (1982), Berg, Dickhaut, and McCabe (1995), Fehr and Gächter (2000), and Camerer (2003, Chapter 2).

[11] There exist a large number of experimental studies on infinitely repeated games with perfect monitoring. See Roth and Murnighan (1978), Murnighan and Roth (1983), Dal Bó (2005), Dal Bó and Fréchette (2010), and others.

[12] See Rabin (1993), Charness and Rabin (2002), Falk, Fehr, and Fischbacher (2003), Dufwenberg and Kirchsteiger (2004), and Falk and Fishbacher (2005). For a survey on related issues, see Sobel (2005).

3. The Model

We investigate a *repeated game* played by players 1 and 2 in a discrete time horizon. The game ends in a finite number of rounds, and the terminal round is randomly determined. The *component game* of this repeated game is given 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 $s \in S$. We assume each player $i's$ payoff has an additively separable form;

$$u_i(s) = v_i(s_i) + w_i(s_j) \text{ for all } s \in S \text{, where } j \neq i.$$

Two random signals, $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$, occur after action choices are made, where Ω_i denotes the set of possible ω_i, $\omega = (\omega_1, \omega_2)$, and $\Omega = \Omega_1 \times \Omega_2$. A signal profile $\omega \in \Omega$ is randomly determined according to the conditional probability function $f(\cdot \mid s) : \Omega \to R_+$. Let $f_i(\omega_i \mid s) \equiv \sum_{\omega_j \in \Omega_j} f(\omega \mid s)$. We assume that $f_i(\omega_i \mid s)$ is independent of s_j; we denote $f_i(\omega_i \mid s_i)$ instead of $f_i(\omega_i \mid s)$ and use $\omega_i \in \Omega_i$ to denote *the signal for player* $i's$ *action.*

We assume monitoring is *imperfect*: at every round $t \in \{1,2,...\}$, player i cannot directly observe the action $s_j(t) \in A_j$ his partner $j \neq i$ selected. He also cannot observe the realized payoff profile $u(s(t)) = (u_1(s(t)), u_2(s(t)))$. Instead, he receives a signal for his partner's action $\omega_j(t) \in \Omega_j$, through which he can imperfectly monitor his partner's choice. We assume monitoring is *private*; player i cannot observe the signal for his own action $\omega_i(t) \in \Omega_i$.[13]

Let $h(t) = (s(\tau), \omega(\tau))_{\tau=1}^{t}$ denote the *history up to round* t. Let us denote by $H = \{h(t) \mid t = 0,1,...\}$ the set of possible histories, where $h(0)$ implies the null history. *The payoff for player* i *per round with the history* $h(t) \in H$ *up to round* t is defined as

[13] We assume for the experiments that the subjects observe the realization of their payoffs after the terminal time.

$$U_i(h(t)) \equiv \frac{\sum_{\tau=1}^{t} u_i(s(\tau))}{t}.$$

Let us specify the component game as a *prisoners' dilemma with symmetry and additive separability*; for each $i \in \{1,2\}$,

$$S_i = \{A, B\},\ v_i(A) = -Y,\ v_i(B) = 0,\ w_i(A) = X + Y, \text{ and}$$

$$w_i(B) = X + Y - Z,$$

where X, Y, and Z are positive integers, and $Z > Y > 0$. Let us call A the *cooperative* action and B the *defective* action. It costs player i Y if he selects a cooperative action choice, but it gives the partner a benefit Z, which is greater than the cost Y. Note the payoff vector (X, X) induced by the cooperative action profile (A, A) is efficient, and is better than the payoff vector $(X+Y-Z, X+Y-Z)$ that is induced by the defective action profile (B, B), which is a dominant strategy profile and the unique Nash equilibrium in the component game.

In the experiment, we use $(X, Y, Z) = (60,10,55)$. If both players cooperate, they achieve the outcome $(X, X) = (60,60)$. If player 1 cooperates and player 2 defects, they achieve the outcome $(X-Z, X+Y) = (5,70)$. If player 1 defects and player 2 cooperates, they achieve the outcome $(X+Y, X-Z) = (70,5)$. If both players defect, they achieve the outcome $(X+Y-Z, X+Y-Z) = (15,15)$. The payoff matrix is described in Figure 1.

[Figure 1]

Let us specify

$$\Omega_i = \{a, b\} \text{ and } f_i(a \mid A) = f_i(b \mid B) = p, \text{ where } 1/2 < p < 1.$$

Let us call a the *good* signal, and b the *bad* signal. The probability index p implies t*he level of monitoring accuracy*; the greater p is, the more accurate the signal. The inequality $p > 1/2$ implies that it is more likely for player i to receive a good signal when his partner selects the cooperative action A rather than the defective action B.

For each history $h(t) \in H$ up to round t, let us denote the frequency of cooperative action choice A by

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

From additive separability, it follows the sum of the payoffs per round with the history $h(t) \in H$ up to round t is given by

$$U_1(h(t)) + U_2(h(t)) = 2[X + \{1 - \rho(h(t))\}(Y - Z)].$$

This implies for given X, Y, and Z, the frequency of cooperative action choice $\rho(h(t))$ uniquely determines the sum of the payoffs per round.

Let $\delta \in (0,1)$ denote the probability of the repeated game continuing at the end of each round when this game has continued up to the previous round $t-1$. Then, the game is terminated at the end of each round $t \geq 1$ with probability $\delta^{t-1}(1-\delta)$. Hence, the expected number of rounds of the repeated game is given by $\sum_{t=1}^{\infty} \delta^{t-1}(1-\delta)t$.

For our experiments, we use $\delta = 0.967$. This probability is sufficiently high. So, it will be beneficial for a self-interested player, who is motivated just by his own monetary payoff, to coordinate with his partner to achieve the cooperative outcomes, even if monitoring is inaccurate.

4. Symmetric Generous Tit-for-Tat Nash Equilibrium

Let $\alpha_i \in [0,1]$ denote a mixed action for player i; he makes the cooperative action choice A with probability α_i. Player $i's$ strategy in the repeated game is defined as $\sigma_i : H \to [0,1]$; he selects A with probability $\alpha_i = \sigma_i(h(t-1))$ in each round t with the history $h(t-1)$ up to round $t-1$. Let Σ_i denote the set of all strategies for player i, $\sigma = (\sigma_1, \sigma_2)$, and $\Sigma = \Sigma_1 \times \Sigma_2$. The expected payoff per period for player i induced by $\sigma \in \Sigma$ when the monitoring accuracy is given by $p \in (0,1)$ is defined by

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

where $E[\cdot \mid \sigma, p]$ denotes the expectation operator. The expected frequency of cooperative action choice A induced by $\sigma \in \Sigma$ when monitoring accuracy is given by $p \in (0,1)$ is defined by

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

From additive separability, it follows

$$U_1(\sigma;p) + U_2(\sigma;p) = 2[X + \{1-\rho(\sigma;p)\}(Y-Z)].$$

Hence, for given X, Y, and Z, the expected frequency $\rho(\sigma;p)$ uniquely determines the sum of the expected payoffs per period $U_1(\sigma;p) + U_2(\sigma;p)$.

A strategy profile $\sigma \in \Sigma$ is said to be a *Nash equilibrium* in the repeated game with monitoring accuracy $p \in (0,1)$ if

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

A strategy profile $\sigma \in \Sigma$ is said to be *symmetric generous tit-for-tat* if there exists $(q, r(a), r(b)) \in [0,1]^3$ such that $r(a) > 0$, $r(a) \neq r(b)$,

$$\sigma_1(h(0)) = \sigma_2(h(0)) = q,$$

and for each $i \in \{1,2\}$, every $t \geq 2$, and every $h(t-1) \in H$,

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

and

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

At round 1, each player makes the cooperative action choice A with probability q. At each round $t \geq 2$, each player i makes the cooperative action choice A with probability

$r(\omega_j)$ when he observes the signal $\omega_j(t-1)=\omega_j$ for his partner's action at the previous round $t-1$. We will write $(q,r(a),r(b))$ instead of σ for any symmetric generous tit-for-tat strategy profile.[14]

Let us define

(1) $$w(p)\equiv\frac{Y}{\delta(2p-1)Z}.$$

Note $0<w(p)\le 1$ if and only if $\delta\ge\frac{Y}{(2p-1)Z}$.

The following proposition demonstrates if a symmetric generous tit-for-tat strategy profile $(q,r(a),r(b))$ is a Nash equilibrium, then the difference in probability of making the cooperative action choices between the good signal and the bad signal, i.e., $r(a)-r(b)$, must be equal to the value $w(p)$ given by (1).

Proposition 1: *A symmetric generous tit-for-tat strategy profile* $(q,r(a),r(b))$ *is a Nash equilibrium in the repeated game with monitoring accuracy* p *if and only if*

(2) $$\delta\ge\frac{Y}{(2p-1)Z},$$

and

(3) $$r(a)-r(b)=w(p).$$

Proof: Selecting $s_i=A$ instead of B costs player i Y at the current round, whereas in the next round he can gain Z from his partner's response with a probability of $pr(a)+(1-p)r(b)$ instead of $(1-p)r(a)+pr(b)$. This holds irrespective of rounds and history because of additive separability. Since he is incentivized to select either A and B at all times, indifference to these action choices must be a necessary and sufficient condition:

---

[14] The term of "generous" implies that it might be the case that $r(b)>0$, i.e., a player selects the cooperative action with positive probability even if he observes the bad signal.

$$-Y + \delta Z\{pr(a) + (1-p)r(b)\} = \delta Z\{(1-p)r(a) - pr(b)\},$$

which is equivalent to

$$r(a) - r(b) = \frac{Y}{\delta(2p-1)Z}.$$

The above equation corresponds to equality (3). Since $r(a) - r(b) \leq 1$, it follows $1 \geq \frac{Y}{\delta(2p-1)Z}$. Thus, inequality (2) must hold.[15]

Q.E.D.

By Proposition 1, we will write $(\tilde{q}, \tilde{r}) \in [0,1]^2$ instead of $(q, r(a), r(b))$ for any symmetric generous tit-for-tat Nash equilibrium, where

$$\tilde{q} \equiv q,\ \ \tilde{r} \equiv r(a), \text{ and } \ r(b) = \tilde{r} - w(p) = \tilde{r} - \frac{Y}{\delta(2p-1)Z}.$$

Since player i is indifferent between the choices A and B at all times, it follows from additive separability that the expected frequency of the cooperative action choice induced by $(\tilde{q}, \tilde{r})$ is given by

$$\rho(\tilde{q}, \tilde{r}; p) = \frac{1}{Z-Y}[\{(1-\delta)\tilde{q} + \delta\tilde{r}\}Z - \frac{p}{2p-1}Y],$$

and the expected payoff per round induced by $(\tilde{q}, \tilde{r})$ is given by

$$U_1(\tilde{q}, \tilde{r}; p) = U_2(\tilde{q}, \tilde{r}; p) = X + \{1 - \rho(\tilde{q}, \tilde{r}; p)\}(Y - Z).$$

Let us define the intensity with which a subject retaliates against his partner as the difference in frequency of the cooperative action choices between the good signal and the bad signal. The greater the difference in frequency of the cooperative action choices between the good signal and the bad signal, the more severely a subject retaliates against his partner. If a self-interested subject plays a symmetric generous tit-for-tat Nash

[15] The generous tit-for-tat Nash equilibrium satisfies belief-freeness in the sense that each player's incentive constraint is irrelevant to his belief about which signal the opponent has observed. Ely and Välimäki (2002) explore the belief-free construction in repeated games with private monitoring.

equilibrium, this difference can be approximated by the difference in probability of making the cooperative action choices between the good signal and the bad signal, i.e., $w(p)$.

It's important to note that $w(p)$ is *decreasing in* p; the less accurate the monitoring technology is, the more severely a self-interested player retaliates against his partner. In order to incentivize the player to select the cooperative action, it's necessary for a player to select the defective action when he observes a bad signal rather than a good signal. When monitoring is less accurate, the player is more likely to receive the bad signal even when the partner has selected the cooperative action. This might cause a *welfare loss*; the player might retaliate even when his partner has actually selected the cooperative action. On the other hand, if monitoring technology is more accurate, a player can better incentivize his partner by using a milder punishment. It could decrease the welfare loss caused by monitoring imperfection.

5. Experimental Design

We conducted computerized experiments on October 5 and 6, 2006, at the University of Tokyo, where subjects were recruited from the undergraduate and graduate schools in all fields.[16] The subjects received points equal to their earned payoffs, which were converted into yen (0.6 yen per point). They received a participation fee of 1500 yen in addition to their earned payoffs.

We specify the prisoners' dilemma with symmetry and additive separability as $(X,Y,Z)=(60,10,55)$. The payoff matrix is described in Figure 1. The probability of the session continuing to the following round is given by $\delta = 0.967$; the probability of the session being terminated is set to be low enough to make sure self-interested players have incentives to cooperate with each other to some extent, even if monitoring is inaccurate.

The sessions are categorized into two types: type 0.9 in which monitoring accuracy is $p=0.9$, and type 0.6 in which monitoring accuracy is $p=0.6$. Type 0.9 obviously has

[16]The experiment was programmed and conducted with the software z-Tree. See Fischbacher (2007).

greater monitoring accuracy than type 0.6.

We conducted four experiments, and each experiment comprised six sessions, which consisted of several rounds. The number of experimental subjects and the number of rounds in each session are summarized in Table 1. Experiment 1 and Experiment 2 were conducted on October 5, 2006. In each experiment, three sessions of type 0.6 were played, followed by three sessions of type 0.9. Experiment 1 consists of three sessions of type 0.6 with 24, 40, and 25 rounds, and three sessions of type 0.9 with 28, 33, and 14 rounds. Experiment 2 contains three sessions of type 0.6 with 20, 23, and 37 rounds, followed by three sessions of type 0.9 with 34, 34, and 19 rounds. On October 6, 2006, two experiments, Experiments 3 and 4, were conducted with three sessions of type 0.9, followed by three sessions of type 0.6. Experiment 3 comprises three sessions of type 0.9 with 38, 21, and 25 rounds, followed by three sessions of type 0.6 with 25, 28, and 29 rounds. Experiment 4 includes three sessions of type 0.9 with 25, 35, and 23 rounds, and three sessions of type 0.6 with 36, 30, and 21 rounds (See Table 1).

Subjects were randomly paired at the beginning of each session; the pairs remained unchanged throughout the session. They were rematched with another partner at the beginning of the following session. Therefore, each subject had 6 different partners throughout the experiment, i.e., one partner per session.

[Table 1]

Each subject was given an experiment manual with instructions about game rules and printed computer screen images in Japanese.[17] These contents were explained by a voice recording. On the computer screen, each subject could always see the structure of the game, the history of his own action choices and the signals for his partner's actions.

It is important to note that the subjects were *not* informed in advance about how many rounds there are in each session. They were only informed that the number of rounds was randomly determined according to the probability $1-\delta = 0.033$ of termination. In order to help the subjects understand this random termination, we showed them Figure A.1 in the

[17] See the supplement of this paper for the translation of the experiment manual and computer screen images into English.

Appendix on the computer screen at the end of each round when the session does not end at that round. When the session ends at a given round, Figure A.2 in the Appendix was shown on the computer screen. These screens were shown to help them understand that one of the 30 cells (numbered 1 to 30) is selected at random, and the session is terminated if and only if the 30th cell (number 30) is selected. When one of other 29 cells (numbers 1 to 29) is selected, the session continues to the next round.

Roth and Murnighan (1978) and Murnighan and Roth (1983) established the experimental design for infinitely repeated games with random termination in which time discounting is replicated by a fixed probability had the game continues at the end of each round. Dal Bó (2005) reports subjects are more successful in implicit collusion in infinitely repeated games with random termination than in finitely repeated games with a known terminal round. The present paper uses a random termination device as a proper replication of time discounting similar to these papers.

6. Experimental Results

Table 2 summarizes the mean frequencies of cooperative action choices in each experiment denoted by

$$\rho(p) \equiv \frac{Num(A;p)}{Num(A;p)+Num(B;p)}$$

where $Num(A;p)$ and $Num(B;p)$ denote the number of times the cooperative action A and the cooperative action B where selected for a given level of accuracy p ($p = 0.9$ or 0.6), respectively. We conduct two analyses. In the first analysis, we use individual actions as observations, and calculate the mean frequencies of cooperative action choices. In the second analysis, we calculate the mean frequencies of cooperative actions for each individual, and use the individual means as observations.

The first part of Table 2 shows the result of the first analysis that uses individual actions as observations. It demonstrates the cooperative action A is chosen more often when signals are more accurate, i.e., $p = 0.9$, than when signals are less accurate, i.e., $p = 0.6$, in all four experiments. On average, the cooperative action was selected 67.2 percent of times when monitoring accuracy is 90 percent across all experiments, while the cooperative action was selected only 35.5 percent of times on average when the monitoring accuracy is low ($p = 0.6$). We conduct a Mann-Whitney rank-sum test. The null hypothesis is there is no difference in the frequencies of the cooperative action choices between $p = 0.9$ and $p = 0.6$. The null hypothesis is rejected at the one percent significance level in all four experiments. The lower part of Table 2 shows the result of the second analysis that uses the means of individual's actions as observations. On average, subjects choose the cooperative action 67 percent of times when monitoring accuracy is 90 percent, while they choose the cooperative action only 35.4 percent of times when monitoring accuracy is low ($p = 0.6$).

This result is consistent with the theoretical prediction that when signals are highly accurate, players could send good signals and effectively coordinate to achieve the cooperative outcome in future periods. We also checked whether the above results hold when we use data only from the first round from sessions, rounds 2 to 14, and rounds 15 and above. We confirmed that the results hold in all segments of the data (See Table A.1 in the Appendix).

[Table 2]

Table 3 and Table 4 show the mean frequencies of cooperative action choices when subjects receive good signals and bad signals, respectively. The frequency of the cooperative action choices contingent on receiving good signals is defined as

$$r(a;p) \equiv \frac{Num(A;a,p)}{Num(A;a,p) + Num(B;a,p)},$$

where $Num(A;a,p)$ and $Num(B;a,p)$ denote the number of times the cooperative action A and the cooperative action B were selected by subjects when they have received good signals, respectively. $r(b;p)$ is similarly defined. Similar to Table 2, we conducted two analyses, first using individual actions as observations, and second using means of individual's action choices as observations. Conducting the second analyses is particularly important in Tables 3 and 4. In the first analyses, the actions of those subjects who happened to have cooperative and uncooperative partners (thus receiving more good and bad signals) become overrepresented in Tables 3 and 4, respectively.

Table 3 shows having received good signals in the previous period, the subjects choose the cooperative action A more often when signals are more accurate ($p = 0.9$) than less accurate ($p = 0.6$). On average, the cooperative action was selected 85.2 percent of times when the monitoring accuracy is 90 percent, and the cooperative action was selected only 43.7 percent of times on average when the monitoring accuracy is low ($p = 0.6$). We conduct a Mann-Whitney rank-sum test with the null hypothesis that there is no difference in the frequencies of cooperative action choices between $p = 0.9$ and $p = 0.6$ when subjects have received good signals. The null hypothesis is rejected in all four sessions at the one percent significance level. We also checked whether the above results hold when we use data only from the first round of sessions, rounds 2 to 14, and rounds 15 and above. We confirmed that the results hold in all fragments (See Table A.2 in the Appendix). The second part of Table 3 shows the subjects select the cooperative action 78.8 percent of times on average if they have received good signals and $p = 0.9$. When $p = 0.6$, the subjects select the cooperative action 42.3 percent of times if they have received good signals.

Table 4 shows the frequencies of cooperative action choices when subjects have received bad signals in the previous periods. Table 4 tells us that even when subjects have received bad signals, they still choose the cooperative action A more often under the higher accuracy of signals ($p = 0.9$) than the lower accuracy of signals ($p = 0.6$). On average, the cooperative action was selected 34.3 percent of times when the monitoring accuracy is 90 percent, while it was selected only 27.2 percent of times when the monitoring accuracy is 60 percent (see upper part of Table 4). The gap between frequencies of cooperate action choices under the high signal-accuracy treatment and the low signal-accuracy treatment is narrower when subjects have received bad signals than good signals. We conduct a Mann-Whitney rank-sum test with the null hypothesis that there is no difference in the frequencies of the cooperative action choices between $p = 0.9$ and $p = 0.6$ when subjects have received bad signals. The null hypothesis is rejected in three out of four experiments at the one percent significance level. This result is consistent with the theoretical prediction when signals are highly accurate ($p = 0.9$), players who have received bad signals from their partners could still coordinate to realize the cooperative outcome in future periods more easily compared with the case in which the level of signal accuracy is low ($p = 0.6$). We also checked whether the above result is maintained when we use data only from the first round of sessions, rounds 2 to 14, and rounds 15 and above. We confirmed that the results hold in all three segments (See Table A.3 in the Appendix).

The second part of Table 4 uses the means of individual data as observations. On average, the subjects choose the cooperative action 44.8 percent of times when they have received the bad signal and $p = 0.9$, while they choose the cooperative action 27.9 percent of times when they have received the bad signal and $p = 0.6$. The reason why the mean frequency of cooperative action is higher for the second analysis than the first data analysis suffers from sample bias. When we use actions as observations, the subjects who have received the bad signals from their partners more frequently become over represented. These individuals are less likely to choose the cooperative actions.

[Table 3]

[Table 4]

Table 5 displays the differences in mean frequencies of cooperate action choices when subjects have received good signals and bad signals, i.e., $r(a; p) - r(b; p)$, for a given level of signal accuracy, p, respectively.

Notice subjects select the cooperative action more often when they have received good signals rather than bad signals regardless of the accuracy levels of signals. This also indicates subjects are more likely to retaliate against their partners when they have received bad signals instead of good signals. This result is consistent with the theoretical prediction that a player incentivizes his partner by rewarding when he has received a good signal, and retaliates against his partner when he has received a bad signal. However, the gap between the frequencies of cooperative action choices when subjects have received good signals and bad signals, $r(a;p)-r(b;p)$, is larger when the level of signal accuracy is higher. Table 5 summarizes the difference between mean frequencies of cooperative actions between good signals and bad signals. When we use each action as an observation, the mean difference is 0.508 for $p=0.9$. When $p=0.6$, the mean difference in signal-contingent frequencies is much smaller (0.165). When we use the mean action of individuals as observations, the gap is narrower for $p=0.9$ (35.2 percent). It is due to the sample bias discussed above.

These experimental results contrast with the prediction of a symmetric generous tit-for-tat Nash Equilibrium. In Section 4, we defined the intensity with which a subject retaliates against his partner as the difference in frequency of the cooperative action choices between the good signal and the bad signal. The greater the difference in frequency of the cooperative action choices between the good signal and the bad signal is, the more severely a subject retaliates against his partner. In a symmetric generous tit-for-tat Nash Equilibrium, the gap between the frequencies of cooperative action choices when subjects have received good signals and bad signals, is decreasing in the level of signal accuracy p. However, our experiments yield opposite results. In the experimental results, this gap is rather increasing in the level of signal accuracy p.

[Table 5]

The latter part of this section compares our experimental results with the theoretical prediction discussed in Section 4. We use all actions as observations for computation. The substance of the results is unchanged if we use the means of individual data as observations.

We assume experimental subjects play a symmetric generous tit-for-tat Nash equilibrium $(\tilde{q},\tilde{r})=(\tilde{q}(p),\tilde{r}(p))$ specified by

$$(\tilde{q}(p),\tilde{r}(p))=(q(p),r(a;p)) \quad \text{if} \quad r(a;p)-w(p)\geq 0,$$

and

$$(\tilde{q}(p),\tilde{r}(p))=(0,w(p)) \quad \text{if} \quad r(a;p)-w(p)<0.$$

$\tilde{q}(p)$ denotes the theoretical prediction of the probability players select the action choice A in round 1. $q(p)$ represents the mean probability the subjects select the action choice

A in round 1 in the experiments. $\tilde{r}(p)$ designates the theoretical prediction of the probability players select the action choice A in any round after round 1 for given p. $r(a;p)$ is the mean probability the subjects select the action choice A in round 2 and afterward in the experiments when subjects have received good signals. $w(p)$ is defined as (1) for the accuracy p, i.e., $w(p) \equiv \dfrac{Y}{\delta(2p-1)Z}$.

Suppose $r(a;p) - w(p) \geq 0$. Then, there exists a symmetric generous tit-for-tat strategy profile $(\tilde{q}, \tilde{r}(a), \tilde{r}(b))$ that is a Nash equilibrium:

$$(\tilde{q}, \tilde{r}(a), \tilde{r}(b)) = (q(p), r(a;p), r(a;p) - w(p)).$$

We shall regard this strategy profile as the theoretical prediction based on the self-interested motives: $(\tilde{q}(p), \tilde{r}(p)) = (\tilde{q}, r(a;p))$.

On the other hand, if $r(a;p) - w(p) < 0$, there exists no such symmetric generous tit-for-tat strategy profile. In this case, we assume $\tilde{r}(p) = w(p)$ and define a symmetric generous tit-for-tat Nash equilibrium as $(\tilde{q}(p), \tilde{r}(p)) = (0, w(p))$. This is the worst symmetric generous tit-for-tat Nash equilibrium.

From Section 4, we can calculate $w(0.9)$, $w(0.6)$, $\rho(\tilde{q}, \tilde{r}; 0.9)$, and $\rho(\tilde{q}, \tilde{r}; 0.6)$ as follows:

$$w(0.9) \approx 0.235,$$

$$w(0.6) \approx 0.94,$$

$$\rho(\tilde{q}, \tilde{r}; 0.9) = \frac{11(0.033\tilde{q} + 0.967\tilde{r})}{9} - \frac{1}{4},$$

$$\rho(\tilde{q}, \tilde{r}; 0.6) = \frac{11(0.033\tilde{q} + 0.967\tilde{r})}{9} - \frac{2}{3}.$$

Table 6 summarizes the implied symmetric generous tit-for-tat Nash equilibria. Note $(\tilde{q}(0.6), \tilde{r}(0.6))$ is specified as the worst symmetric generous tit-for-tat Nash equilibrium because $w(0.6) > r(a;0.6)$.

[Table 6]

From Tables 2 and 6, it follows

$$\rho(0.9) < \rho(\tilde{q},\tilde{r};0.9) \text{ and } \rho(0.6) < \rho(\tilde{q},\tilde{r};0.6),$$

which implies irrespective of monitoring accuracy, a subject is less likely to select the cooperative action than the specified symmetric generous tit-for-tat Nash equilibrium predicts. From Tables 5 and 6, it follows:

$$r(a;0.9) - r(b;0.9) > w(0.9) .$$

When $p = 0.9$, i.e., when the monitoring technology is highly accurate, the subject tends to retaliate against his partner more severely than this equilibrium predicts. This tendency increases the welfare loss caused by the monitoring imperfection.

From Tables 5 and 6, it follows:

$$r(a;0.6) - r(b;0.6) < w(0.6) .$$

When $p = 0.6$, i.e., the monitoring technology is less accurate, the subject tends to retaliate against his partner less severely than the symmetric generous tit-for-tat Nash equilibrium predicts. Hence, when the monitoring technology is less accurate, the subjects are underutilizing retaliation to incentivize their partners to select the cooperative action. This seems to be the vital reason for welfare loss when the monitoring technology is less accurate.

7. Behavioral Models

This section provides a behavioral model. We incorporate bounded rationality to the model discussed in Section 3, and propose an alternative model that can explain behavioral deviations from the generous tit-for-tat Nash equilibria with purely self-interested players. Under imperfect private monitoring, it’s difficult for people to behave rationally all the time. Moreover, players might be motivated not only by pure self-interest but also by social preferences. Hence, we assume 1) players select actions randomly with some probability, 2) players suffer from a psychological cost of guilt if they select a defective action after having observed a good signal, and 3) players bear a psychological cost of resentment if they select a cooperative action after having observed a bad signal.

To be precise, a player behaves in the repeated games according to the manner that:

(i) With probability $\varepsilon = \varepsilon(p) \in [0, \frac{1}{2}]$, he selects A.

(ii) With probability ε, he selects B.

(iii) With the remaining probability $1-2\varepsilon$, he selects an action that maximizes his long-term expected utility that incorporates the following psychological costs:

a) he suffers from a psychological cost of guilt $c(a) = c(a; p) \geq 0$ if he chooses the defective action B after having observed a good signal a.

b) he suffers from a psychological cost of resentment $c(b) = c(b; p) \geq 0$ if he takes the cooperative action A after having observed a bad signal b.

We replace the assumption of pure self-interests with behavioral assumptions specified above. Recall when his partner follows $(q, r(a), r(b))$, a player loses Y in the current round if he selects the cooperative action A, but in the next round he gains Z from his partner's cooperative response with probability $pr(a) + (1-p)r(b)$ instead of $(1-p)r(a) + pr(b)$. In addition, he can save the psychological cost of guilt $c(a)$ if he has observed the good signal a and takes the cooperative action A, while he suffers from the psychological cost of resentment $c(b)$ when he has observed the bad signal b and chooses the cooperative action A.

Associated with the behavioral model described by $(\varepsilon, c(a), c(b))$, we define a solution concept named quasi-Nash equilibrium. In order to focus on the responses to signals, we weaken the standard Nash equilibrium concept by eliminating the incentive requirement at the first round. A symmetric generous tit-for-tat strategy profile $(q, r(a), r(b))$ is said to be a *quasi-Nash equilibrium* associated with behavioral model $(\varepsilon, c(a), c(b))$ if

$$\min[r(a), 1-r(a), r(b), 1-r(b)] \geq \varepsilon,$$

(4) $$[-Y + \delta Z\{pr(a) - (1-p)r(b)\} < \delta Z\{(1-p)r(a) - pr(b)\} - c(a)]$$
$$\Rightarrow [r(a) = \varepsilon],$$

(5) $$[-Y + \delta Z\{pr(a) - (1-p)r(b)\} > \delta Z\{(1-p)r(a) - pr(b)\} - c(a)]$$

$\Rightarrow [1 - r(a) = \varepsilon]$,

(6) $[-Y + \delta Z\{pr(a) - (1-p)r(b)\} - c(b) < \delta Z\{(1-p)r(a) - pr(b)\}]$

$\Rightarrow [r(b) = \varepsilon]$,

and

(7) $[-Y + \delta Z\{pr(a) - (1-p)r(b)\} - c(b) > \delta Z\{(1-p)r(a) - pr(b)\}]$

$\Rightarrow [1 - r(b) = \varepsilon]$.

The equation (4) describes the case in which a player has received a good signal from his partner, and attains higher utility if he selects the defective action. Then, the probability that the player takes the cooperative action is at the minimum level, i.e., $r(a) = \varepsilon$. The equation (5) explains the case in which a player has received a good signal from his partner, and attains higher utility if he chooses the cooperative action. Then, the probability that the player selects the defective action is at the minimum level, i.e., $1 - r(a) = \varepsilon$. The equation (6) corresponds to the case in which a player has receives a bad signal from his partner, and attains higher utility if he takes the cooperative action. Then, the probability the player selects the defective action is at the minimum level, i.e., $r(b) = \varepsilon$. The equation (7) relates to the case in which a player has receives a bad signal from his partner, and attains higher utility if he choose the cooperative action. Then, the probability the player selects the defective action is at the minimum level, i.e., $1 - r(b) = \varepsilon$.

The following theorem characterizes quasi-Nash equilibrium.

Theorem 2: *Suppose* $r(a) - r(b) \neq w(p)$ *and* $1 - r(a) \neq r(b)$. *Then, a symmetric generous tit-for-tat strategy profile* $(q, r(a), r(b))$ *is a quasi-Nash equilibrium associated with a behavioral model* $(\varepsilon, c(a), c(b))$ *if and only if either*

$r(a) - r(b) > w(p)$,

$\varepsilon = 1 - r(a)$ *and*

$c(b) = \delta Z(2p-1)\{r(a) - r(b)\} - Y = \delta Z(2p-1)[\{r(a) - r(b)\} - w(p)]$,

or

$r(a) - r(b) < w(p)$,

$\varepsilon = r(b)$ *and*

$c(a) = Y - \delta Z(2p-1)\{r(a) - r(b)\} = \delta Z(2p-1)[w(p) - \{r(a) - r(b)\}]$.

Proof: Note if

$$-Y + \delta Z\{pr(a) - (1-p)r(b)\} > \delta Z\{(1-p)r(a) - pr(b)\},$$

that is,

$$r(a) - r(b) > w(p),$$

then the left hand side of (2) holds, i.e., $\varepsilon = 1 - r(a)$, which along with $1 - r(a) \neq 1 - r(b) > \varepsilon$, (3), and (4) implies that either $r(b) = \varepsilon$ or

$$-Y + \delta Z\{pr(a) - (1-p)r(b)\} - c(b) = \delta Z\{(1-p)r(a) - pr(b)\}.$$

Since $\varepsilon = 1 - r(a) \neq r(b)$, it follows

$$c(b) = \delta Z(2p-1)\{r(a) - r(b)\} - Y.$$

Note if

$$-Y + \delta Z\{pr(a) - (1-p)r(b)\} < \delta Z\{(1-p)r(a) - pr(b)\},$$

that is,

$$r(a) - r(b) < w(p),$$

then the left hand side of (3) holds, i.e., $\varepsilon = r(b)$, which along with $r(b) \neq r(a) > \varepsilon$, (1), and (2) implies that either $1 - r(a) = \varepsilon$ or

$$-Y + \delta Z\{pr(a) - (1-p)r(b)\} = \delta Z\{(1-p)r(a) - pr(b)\} - c(a).$$

Since $\varepsilon = r(b) \neq 1 - r(a)$, it follows

$$c(a) = Y - \delta Z(2p-1)\{r(a) - r(b)\}.$$

Q.E.D.

By using individual actions as observations, we can specify symmetric generous tit-for-tat strategy profiles $(q, r(a), r(b))$ as follows:

$$(r(a), r(b)) = (0.437, 0.272) \quad \text{if} \quad p = 0.6.$$

$(r(a), r(b)) = (0.852, 0.344)$ if $p = 0.9$.

Let us specify the parameters in the behavioral model as follows:

$\varepsilon(0.6) = 0.272$, $\varepsilon(0.9) = 0.148$, $c(a; 0.6) = 8.245$, and $c(b; 0.9) = 11.614$.

Note the above symmetric generous tit-for-tat strategy profile, which is compatible with the experimental data, is a quasi-Nash equilibrium in the behavioral model associated with the parameters specified for each $p \in \{0.9, 0.6\}$. Hence, the specified parameters are consistent with the experimental data. In the case of $p = 0.9$ (high accuracy), the presence of psychological cost of resentment $c(b)$ lowers player's utility if he selects the cooperative action has after having received a bad signal. As a result, the difference in the frequencies of cooperative action choice between good and bad signals widens in equilibrium. In the case of $p = 0.6$ (low accuracy), the presence of psychological cost of guilt $c(a)$ lowers player's utility for selecting the defective action after having received a good signal. Therefore, the behavioral model predicts for the case of $p = 0.6$, the difference in the frequencies of cooperative action choices between good and bad signals is narrower in equilibrium.

By utilizing the means of individual subject's actions as observations, we can specify symmetric generous tit-for-tat strategy profiles $(q, r(a), r(b))$ as follows:

$(r(a), r(b)) = (0.423, 0.279)$ if $p = 0.6$.

$(r(a), r(b)) = (0.788, 0.448)$ if $p = 0.9$.

Let us specify the parameters in the behavioral model as follows:

$\varepsilon(0.6) = 0.279$, $\varepsilon(0.9) = 0.212$, $c(a; 0.6) = 8.469$, and

$c(b; 0.9) = 4.467$.

Even when we use the means of individual's actions as observations, our results hold.

We have developed an alternative behavioral model and tested it using the experimental results for $p = 0.6$ and $p = 0.9$. We have confirmed the derived parameters in the behavioral model are consistent with the experimental results. Below, we extend the behavioral model to more general settings for any monitoring accuracy $p \in (1/2, 1)$.

The experimental results indicate 1) subjects are more likely to select the cooperative action when they have observed good signals than bad signals. 2) Irrespective of which signal they have observed, subjects select the cooperative action a more often when the signal is more accurate. 3) Also the more accurate the signal technology is, the greater the difference in frequency of cooperative action choices are between good and bad signals.

Based on these indications, let us fix an arbitrary accuracy-contingent symmetric generous tit-for-tat strategy profile given by $((q(p), r(a;p), r(b;p)))_{p\in(\frac{1}{2},1)}$, where we assume $r(a;p) > r(b;p)$, and $r(a;p)$, $r(b;p)$, and $r(a;p) - r(b;p)$ are all increasing and continuous in $p \in (\frac{1}{2},1)$. For technical convenience, and for further relevancy to the experimental results, we assume $r(a;p)$, $r(b;p)$, $1-r(a;p)$, and $1-r(b;p)$ are all different values almost everywhere. This is a reasonable assumption. Note there exists a critical level of accuracy $\hat{p} \in [\frac{1}{2},1]$ such that

$$r(a;p) - r(b;p) > w(p) \quad \text{if} \quad p > \hat{p},$$

and

$$r(a;p) - r(b;p) < w(p) \quad \text{if} \quad p < \hat{p}.$$

Consider any accuracy-contingent behavioral model given by

$$((\varepsilon(p), c(a;p), c(b;p)))_{p\in(\frac{1}{2},1)},$$

where we assume $\varepsilon(p)$, $c(a;p)$, and $c(b;p)$ are all continuous in p, and for every $p \in (\frac{1}{2},1)$,

$$\varepsilon(p) \leq \min[r(a;p), 1-r(a;p), r(b;p), 1-r(b;p)].$$

The following theorem characterizes the accuracy-contingent behavioral model $((\varepsilon(p), c(a;p), c(b;p)))_{p\in(\frac{1}{2},1)}$ that is compatible with the fixed accuracy-contingent symmetric generous tit-for-tat strategy profile $((q(p), r(a;p), r(b;p)))_{p\in(\frac{1}{2},1)}$ in terms of quasi-Nash equilibrium.

Theorem 3: *For each* $p \in (\frac{1}{2}, 1)$ *except* $\hat{p}$, *the symmetric generous tit-for-tat strategy profile* $(q(p), r(a;p), r(b;p))$ *is a quasi-Nash equilibrium associated with* $(\varepsilon(p), c(a;p), c(b;p))$ *if and only if either*

$$p > \hat{p},$$

$$\varepsilon(p) = 1 - r(a;p), \textit{ and}$$

$$c(b;p) = \delta Z(2p-1)[\{r(a) - r(b)\} - w(p)],$$

or

$$p < \hat{p},$$

$$\varepsilon(p) = r(b;p), \textit{ and}$$

$$c(a;p) = \delta Z(2p-1)[w(p) - \{r(a) - r(b)\}].$$

Proof: Note that $p > \hat{p}$ implies $r(a;p) - r(b;p) > w(p)$, and that $p < \hat{p}$ implies $r(a;p) - r(b;p) < w(p)$. Hence, we can prove this theorem straightforwardly from Theorem 2.

Q.E.D.

Theorem 3 implies that:

1) For an accuracy level greater than the critical level, i.e., $p > \hat{p}$, the probability of naïveté $\varepsilon(p)$ is decreasing, and the psychological cost of resentment $c(b;p)$ is increasing, in p. The psychological cost of guilt $c(a;p)$ is irrelevant.

2) For an accuracy level lower than the critical level, i.e., $p < \hat{p}$, the probability of naïveté $\varepsilon(p)$ is increasing, and the psychological cost of guilt $c(a;p)$ is decreasing, in p. The psychological cost of resentment $c(b;p)$ is irrelevant.

3) With the critical level of accuracy $p = \hat{p}$, the psychological costs are set equal to zero, i.e.,

$$c(a;\hat{p}) = c(b;\hat{p}) = 0,$$

while the probability of naïveté $\varepsilon(p)$ is maximum at $p = \hat{p}$.

Theorem 3 demonstrates the following predictions on players' behavioral motives. The probability of naiveté is single-peaked across accuracies with the peak at the critical level $\hat{p}$. A player is most naïve at the critical level $\hat{p}$. He becomes more sophisticated as the accuracy level moves further away from the critical level $\hat{p}$. At the same time, he is more motivated by social preferences as the accuracy level moves further away from the critical level $\hat{p}$. This implies there is a trade-off between naiveté and social preferences: when a player is making naïve decisions, he is less motivated by social preferences, i.e., reciprocity. Likewise, when a player has strong reciprocal concerns, he makes few random decisions.

Finally, fix an accuracy-contingent behavioral model $((\varepsilon(p), c(a;p), c(b;p)))_{p\in(\frac{1}{2},1)}$ arbitrarily, where we assume there exists a critical level $\hat{p} \in [\frac{1}{2},1)$ that satisfies the above-mentioned properties 1), 2), and 3). We characterize the associated accuracy-contingent quasi-Nash equilibrium $(r(a;p), r(b;p)))_{p\in(\frac{1}{2},1)}$ in the following manner.

Note from Theorem 2 that for any accuracy level $p \in (\frac{1}{2},1)$ and any behavioral model $(\varepsilon, c(a), c(b))$, there may exist two (and only two) quasi-Nash equilibria,[18] $(r(a), r(b))$ and $(\bar{r}(a), \bar{r}(b))$, where the psychological cost of guilt $c(a)$ is relevant to $(r(a), r(b))$, i.e.,

$$r(a) = \varepsilon + w(p) - \frac{c(a)}{\delta Z(2p-1)} \quad \text{and} \quad r(b) = \varepsilon \, ,$$

while the psychological cost of resentment $c(b)$ is relevant to $(\bar{r}(a), \bar{r}(b))$, i.e.,

$$\bar{r}(a) = 1 - \varepsilon \quad \text{and} \quad \bar{r}(a) = 1 - \varepsilon - w(p) - \frac{c(b)}{\delta Z(2p-1)} .$$ [19]

Note that

---

[18] Since q is irrelevant to quasi-Nash equilibria, we write $(r(a), r(b))$ instead of $(q, r(a), r(b))$.

[19] There may exist another equilibrium that is not generous tit-for-tat, such that $r(a) = r(b) = \varepsilon$, i.e., sophisticated players have no incentive to select the defective action, irrespective of the observed signal.

$$r(a)-r(b)=w(p)-\frac{c(a)}{\delta Z(2p-1)}\leq w(p),$$

and

$$\bar{r}(a)-\bar{r}(b)=w(p)+\frac{c(b)}{\delta Z(2p-1)}\geq w(p),$$

where

$$r(a)-r(b)=\bar{r}(a)-\bar{r}(b)=w(p) \quad \text{if} \quad p=\hat{p}.$$ [20]

Consider the behavioral model that corresponds to the experimental data in the case of $p=0.9$, i.e., $\varepsilon=0.148$ and $c(b)=11.614$, where we let $c(a)=0$ arbitrarily. Clearly, $(\bar{r}(a),\bar{r}(b))=(0.852,0.344)$ corresponds to the experimental data. There exists another quasi-Nash equilibrium to which the psychological cost of guilt $c(a)$ is relevant, i.e., $(r(a),r(b))=(0.383,0.148)$, which is much less collusive than $(\bar{r}(a),\bar{r}(b))=(0.852,0.344)$.

If we require $(r(a;p),r(b;p)))_{p\in(1/2,1)}$ to satisfy the experimental data of the intensity of retaliation $r(a;p)-r(b;p)$ being decreasing in p, the accuracy-contingent behavioral model can uniquely determine $(r(a;p),r(b;p)))_{p\in(1/2,1)}$ as the associated accuracy-contingent quasi-Nash equilibrium as follows.

Theorem 4: *Consider an arbitrary accuracy-contingent behavioral model* $((\varepsilon(p),c(a;p),c(b;p)))_{p\in(1/2,1)}$ *and an arbitrary accuracy-contingent quasi-Nash equilibrium* $(r(a;p),r(b;p)))_{p\in(1/2,1)}$. *Suppose there exists* $\hat{p}\in[1/2,1)$ *that satisfies the above properties 1), 2), and 3), and that* $r(a;p)-r(b;p)$ *is increasing in* p. *Then, for every* $p\in(1/2,1)$,

$$r(a;p)=1-\varepsilon(p) \quad \textit{and} \quad r(b;p)=1-\varepsilon(p)-w(p)-\frac{c(b;p)}{\delta Z(2p-1)} \quad \textit{if} \quad p>\hat{p},$$

[20] There exists no other symmetric generous tit-for-tat quasi-Nash equilibrium.

and

$$r(a;p) = \varepsilon(p) + w(p) - \frac{c(a;p)}{\delta Z(2p-1)} \quad \text{and} \quad r(b;p) = \varepsilon(p) \qquad \text{if} \quad p < \hat{p}.$$

Proof: We show below that whenever $p > \hat{p}$, the psychological cost of resentment is relevant to $(r(a;p), r(b;p))$, while whenever $p < \hat{p}$, the psychological of guilt is relevant to $(r(a;p), r(b;p))$.

Suppose $p > \hat{p}$, and the psychological cost of guilt is relevant. Then,

$$r(a;p) - r(b;p) = w(p) - \frac{c(a)}{\delta Z(2p-1)} < w(p) < w(\hat{p}) = r(a;\hat{p}) - r(b;\hat{p}),$$

which contradicts the pattern of experimental data, that is $r(a;p) - r(b;p)$ is increasing in p. Next, suppose $p < \hat{p}$, but the psychological cost of resentment is relevant. Then,

$$r(a;p) - r(b;p) = w(p) + \frac{c(b)}{\delta Z(2p-1)} > w(p) > w(\hat{p}) = r(a;\hat{p}) - r(b;\hat{p}),$$

which again contradicts the experimental data that $r(a;p) - r(b;p)$ is increasing in p. Hence, we have proved this theorem.

Q.E.D.

Theorem 4 implies the generous tit-for-tat strategy profile derived from the experimental data is the unique plausible quasi-Nash equilibrium associated with the underlying behavioral model. This theoretical finding proves our behavioral approach has strong predictive power in terms of both describing strategies and clarifying motives.

8. Conclusion

We proposed a behavioral approach to infinitely repeated games with imperfect private monitoring by conducting laboratory experiment to test the standard game theory, examining motives and strategies behind the behavioral deviations from theoretical predictions, and constructing a behavioral model which is consistent with the behavioral patterns observed in the experiment. We then showed there exists a unique quasi-Nash equilibrium that is consistent with the behavioral patterns.

We conducted experiments with varying monitoring accuracy and showed the experimental results contradict the prediction of standard game theory. In contrast to the theoretical prediction, our subjects retaliate more severely when monitoring is more accurate, making a welfare loss severer than the theory predicts. In the case of low monitoring accuracy, subjects underutilize punishment, making it hard for them to incentivize their partners to select the cooperative action.

Having explored the behavioral patterns from the experimental data, we provided an alternative behavioral model. The proposed model is behavioral in two important ways. First, we incorporate bounded rationality, namely naiveté and social preferences. We assume a player makes a random choice with some probability, and with the remaining probability, he is sophisticated but is motivated by social preferences. A player suffers from the psychological cost of guilt if he selects the defective action after having received a good signal. On the other hand, a player suffers from the psychological cost of resentment if he selects the cooperative action after having received a bad signal. Second, our model is context dependent in that the degrees of naiveté and reciprocal concerns are dependent on the levels of monitoring accuracy. We take into account the fact that people’s motives and perceptions are often context dependent. We showed that the experimental data is consistent with the substantial part of the behavioral model, and that the strategy profile derived from the experimental data is the unique quasi-Nash equilibrium associated with the derived behavioral model.

The accuracy-contingent behavioral model also suggests the structure of players’ motives such as naiveté, social preferences, and pure self-interests. The probability of naiveté is single-peaked across monitoring accuracies. There is a trade-off between naiveté and social preferences in that people make random choices when they do not have strong reciprocal concerns. Conversely, when people are concerned about reciprocity, they make fewer random choices.

Our behavioral approach has strong predictive power not just in clarifying motives as a behavioral model but also in describing a strategy as a unique equilibrium. By using experimental data, we restrict the candidates of equilibria out of numerous equilibria that exist in repeated games with imperfect private monitoring. We then find there exists a unique quasi-Nash equilibrium that is consistent with the behavioral patterns.

It is important to conduct follow-up experiments, change contexts, and attempt to develop behavioral models that better fit various contexts. It might be necessary to replace generous tit-for-tat strategies with a wider class of strategies, where action choices are permitted to be contingent not on the last period observations but on the entire history of observations. The theory of repeated games explore the concept of review strategy, in which players utilize the relative frequency of good signal observations in the history of plays in order to decide whether to retaliate against the opponent now or watch his state for a while. As Rader (1986), Matsushima (2004), and Sugaya (2012) show the concept of review strategy plays a powerful role in facilitating collusion under imperfect private monitoring. Hence, the important question on the line of our behavioral approach would be to study what kinds of behavioral motives will encourage players to make action choices that depend of a history of behaviors.

This paper contributes to the field of repeated games with private monitoring by introducing behavioral approach, and offering an alternative way to refine strategies and equilibria. As far as we know, this paper is the first attempt to propose a behavioral model in infinitely repeated games. This paper also makes an important contribution by offering a solution to a fundamental problem in repeated games. Repeated games have numerous equilibria, and identifying strategies and equilibria has been considered a difficult task. However, by analyzing behavioral patterns in the experiments, we are able to refine a unique quasi-Nash equilibrium that is consistent with behavioral patterns. We also contribute to the social preference literature by extending its idea to the area of repeated games with private monitoring. We believe an accumulation of both experimental and theoretical studies and the interactions of the two fields are necessary to advance behavioral approach to infinitely repeated games with private monitoring, and other areas in game theory.

References

Abreu, Dilip. 1988. "On the Theory of Infinitely Repeated Games with Discounting." *Econometrica* 56(2): 383-396.

Abreu, Dilip, David Pearce, and Ennio Stacchetti. 1990. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica* 58(5): 1041-1063.

Aoyagi, Masaki, and Guillaume Fréchette. 2009. "Collusion as Public Monitoring Becomes Noisy: Experimental Evidence." *Journal of Economic Theory* 114: 1135-1165.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10(1): 122-142.

Camerer, Colin F. 2003. *Behavioral Game Theory.* Princeton University Press.

Charness, Gary, and Martin Dufwenberg. 2006. "Promises and Partnership." *Econometrica* 74(6): 1579-1601.

Charness, Gary, and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117(3): 817-869.

Dal Bó, Pedro. 2005. "Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games." *American Economic Review* 95(5): 1591–1604.

Dal Bó, Pedro, and Guillaume Fréchette. 2010. "Strategy Choice in the Infinitely Repeated Prisoners Dilemma." mimeo.

Dufwenberg, Martin, and George Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47(2): 268-298.

Duffy, John, and Félix Muñoz-García. 2012. "Patience or Fairness? Analyzing Social Preferences in Repeated Games." Games 3(1): 56-77.

Ely, Jeffrey, and Juuso Välimäki. 2002. "A Robust Folk Theorem for the Prisoner's Dilemma." *Journal of Economic Theory* 102(1): 84–105.

Engle-Warnick, Jim, and Robert Slonim. 2006. "Inferring Repeated-Game Strategies from Actions: Evidence from Repeated Trust Game Experiments." *Economic Theory* 28(3): 603–632.

Falk, Armin, Ernst Fehr, and Urs Fischbacher. 2003 "On the Nature of Fair Behavior." *Economic Inquiry* 41(1), 20-26.

Falk, Armin, and Urs Fischbacher. 2005 "Modeling Fairness and Reciprocity," in Moral Sentiments and Material Interests: *The Foundations of Cooperation in Economic Life*. H. Gintis, S. Bowles, R. Boyd, and E. Fehr, eds. Cambridge: MIT Press.

Fehr, Ernst, and Simon Gächter. 2000 "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14(3): 159-181.

Feinberg, Robert, and Christopher Snyder. 2002. "Collusion with Secret Price Cuts: An Experimental Investigation." *Economic Bulletin* 3(6): 1–11.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments."

Experimental Economics 10(2): 171-178.
Fudenberg, Drew, and Jean Tirole. 1993. *Game Theory.* MIT Press.
Geanakoplos, John, David Pearce, and Ennio Stacchetti. 1989. "Psychological Games and Sequential Rationality." *Games and Economic Behavior* 1(1): 60-79.
Green, Edward, and Robert Porter. 1984. "Noncooperative Collusion under Imperfect Price Information." *Econometrica* 52(1): 87–100.
Güth, Werner, Rolf Schmittberger, and Bernd Schwarze. 1982. "An Experimental Analysis of Ultimatum Bargaining." *Journal of Economic Behavior and Organization* 3(4): 367-88.
Holcomb, James, and Paul Nelson. 1997. "The Role of Monitoring in Duopoly Market Outcomes." *Journal of Socio-Economics* 26(1): 79–93.
Kandori, Michihiro. 2002. "Introduction to Repeated Games with Private Monitoring." *Journal of Economic Theory* 102(1): 1-15.
Mailath, George J., and Larry Samuelson. 2006. *Repeated Games and Reputations: Long-Run Relationships*: Oxford University Press.
Mason, Charles F., and Owen R. Phillips. 2002. "In Support of Trigger Strategies: Experimental Evidence from Two-Person Noncooperative Games." *Journal of Economics and Management Strategy* 11(4), 685–716.
Matsushima, Hitoshi. 1991a. "On the theory of repeated games with private information : Part I: anti-folk theorem without communication." *Economics Letters* 35: 253-256.
Matsushima, Hitoshi. 1991b. "On the theory of repeated games with private information: Part II: revelation through communication." *Economics Letters* 35: 257-261.
Matsushima, Hitoshi. 2004. "Repeated Games with Private Monitoring: Two Players." *Econometrica* 72(3): 823–852.
Matsushima, Hitoshi. 2012. "Interlinkage and Generous Tit-For-Tat Strategy." forthcoming in *Japanese Economic Review*.
Matsushima, Hitoshi, and Tomohisa Toyama. 2010. "Accuracy and Retaliation in Repeated Games with Private Monitoring: Theory and Experiments." CIRJE-F-795, University of Tokyo.
Murnighan, J. Keith, and Alvin Roth. 1983. "Expecting Continued Play in Prisoner's Dilemma Games." *Journal of Conflict Resolution* 27(2): 279–300.
Nowak, Martin A., and Karl Sigmund. 1992. "Tit for Tat in Heterogeneous Populations." *Nature* 355: 250–253.
Osborne, Martin J., and Ariel Rubinstein. 1994. *A Course in Game Theory*: MIT Press.
Piccione, Michele. 2002. "The Repeated Prisoners' Dilemma with Imperfect Private Monitoring." *Journal of Economic Theory* 102(1): 70–83.
Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83(5): 1281-1302.
Radner, Roy. 1986. "Repeated Partnership Games with Imperfect Monitoring and No Discounting." *Review of Economic Studies* 53(1): 43-57.
Roth, Alvin, and J. Keith Murnighan. 1978. "Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma." *Journal of Mathematical Psychology* 17: 189–198.
Sekiguchi, Tadashi. 1997. "Efficiency in Repeated Prisoners' Dilemma with Private Monitoring." *Journal of Economic Theory* 76(2): 345–361.

Sobel, Joel. 2005. “Interdependent Preferences and Reciprocity.” *Journal of Economic Literature* 43(2): 392-436.

Sugaya, Takuo. 2012. “The Folk Theorem in Repeated Games with Private Monitoring.” Ph.D. Thesis, Princeton University.

Takahashi, Satoru. 2010. “Community Enforcement When Players Observe Partners’ Past Play.” *Journal of Economic Theory* 145(1): 42–64.

Figure 1:

Prisoners' dilemma with Symmetry and Additive Separability

$(X, Y, Z) = (60, 10, 55)$

	A	B
A	60 60	5 70
B	70 5	15 15

Table 1: Experimental Design

Experiment	Subjects	Type (rounds per session)
Experiment 1 (October 5, 2006)	28	0.6 (24, 40, 25), 0.9 (28, 33, 14)
Experiment 2 (October 5, 2006)	24	0.6 (20, 23, 37), 0.9 (34, 34, 19)
Experiment 3 (October 6, 2006)	28	0.9 (38, 21, 25), 0.6 (25, 28, 29)
Experiment 4 (October 6, 2006)	28	0.9 (25, 35, 23), 0.6 (36, 30, 21)

Table 2: Mean frequencies of Cooperative Action Choices, $\rho(p)$

1) All actions	p=0.9			p=0.6	
Experiment 1	0.748	***	(2,100)	0.401	(2,492)
Experiment 2	0.690	***	(2,088)	0.422	(1,920)
Experiment 3	0.722	***	(2,352)	0.420	(2,296)
Experiment 4	0.538	***	(2,324)	0.193	(2,436)
Total	0.672	***	(8,864)	0.355	(9,144)

2) Means of individuals' actions	p=0.9			p=0.6	
Experiment 1	0.740	***	(28)	0.397	(28)
Experiment 2	0.687	***	(24)	0.417	(24)
Experiment 3	0.718	***	(28)	0.419	(28)
Experiment 4	0.537	***	(28)	0.190	(28)
Total	0.670	***	(108)	0.354	(108)

1) Numbers of observations are in parentheses.

2) *** indicates the mean frequencies of cooperative action choices when p=0.9 are significantly higher than when p=0.6 at the 1 percent significance level by the Mann-Whitney rank-sum test.

Table 3: Mean Frequencies of Cooperative Actions Choices on Good Signals, $r(a;p)$

1) All actions	p=0.9			p=0.6	
Experiment 1	0.887	***	(1,425)	0.480	(1,200)
Experiment 2	0.865	***	(1,315)	0.520	(924)
Experiment 3	0.864	***	(1,539)	0.510	(1,090)
Experiment 4	0.777	***	(1,174)	0.237	(1,051)
Total	0.852	***	(5,453)	0.437	(4,265)
2) Means of individuals' actions	p=0.9			p=0.6	
Experiment 1	0.858	***	(28)	0.474	(28)
Experiment 2	0.813	***	(24)	0.498	(24)
Experiment 3	0.820	***	(28)	0.498	(28)
Experiment 4	0.707	***	(28)	0.235	(28)
Total	0.788	***	(108)	0.423	(108)

1) Numbers of observations are in parentheses.

2) *** indicates the mean frequencies of cooperative action choices when p=0.9 is significantly higher than when p=0.6 at the 1 percent level by the Mann-Whitney rank-sum test.

Table 4: Mean Frequencies of Cooperative Action Choices on Bad Signals, $r(b; p)$

1) All actions	p=0.9			p=0.6	
Experiment 1	0.386	***	(591)	0.315	(1,208)
Experiment 2	0.351		(701)	0.313	(924)
Experiment 3	0.407	***	(729)	0.332	(1,122)
Experiment 4	0.272	***	(1,066)	0.152	(1,301)
Total	0.344	***	(3,087)	0.272	(4,555)
2) Means of individuals' actions	p=0.9			p=0.6	
Experiment 1	0.525	***	(28)	0.320	(28)
Experiment 2	0.438	*	(24)	0.320	(24)
Experiment 3	0.497	***	(28)	0.330	(28)
Experiment 4	0.330	***	(28)	0.154	(28)
Total	0.448	***	(108)	0.279	(108)

1) Numbers of observations are in parentheses.

2) *** indicates the mean frequencies of cooperative action choices when p=0.9 is significantly higher than when p=0.6 at the 1 percent level by the Mann-Whitney rank-sum test.

Table 5: Differences in Mean Frequencies of Cooperative Action Choices between Good Signals and Bad Signals, $r(a;p)-r(b;p)$

1) All actions	p=0.9			p=0.6	
Experiment 1	0.501	***	(2,016)	0.166	(2,408)
Experiment 2	0.514	***	(2,016)	0.208	(1,848)
Experiment 3	0.457	***	(2,268)	0.178	(2,212)
Experiment 4	0.506	***	(2,240)	0.084	(2,352)
Total	0.508	***	(8,540)	0.165	(8,820)
2) Means of individuals' actions	p=0.9			p=0.6	
Experiment 1	0.334	***	(28)	0.154	(28)
Experiment 2	0.375	***	(24)	0.178	(24)
Experiment 3	0.323	**	(28)	0.167	(28)
Experiment 4	0.377	***	(28)	0.081	(28)
Total	0.352	***	(108)	0.144	(108)

1) Numbers of observations are in parentheses.

2) *** indicates the mean frequencies of cooperative action choices when p=0.9 is significantly higher than when p=0.6 at the 1 percent level by the Mann-Whitney rank-sum test.

3) ** indicates the mean frequencies of cooperative action choices when p=0.9 is significantly higher than when p=0.6 at the 5 percent level by the Mann-Whitney rank-sum test.

Table 6: Symmetric Generous Tit-for-Tat Nash Equilibria $(\tilde{q}(p), \tilde{r}(p))$

	$p = 0.9$	$p = 0.6$
$\tilde{q}(p)$	0.781	0.000
$\tilde{r}(p)$	0.852	0.940
$w(p)$	0.235	0.940
$\rho(\tilde{q}(p), \tilde{r}(p); p)$	0.788	0.444

Figure A.1:

Random Termination Device

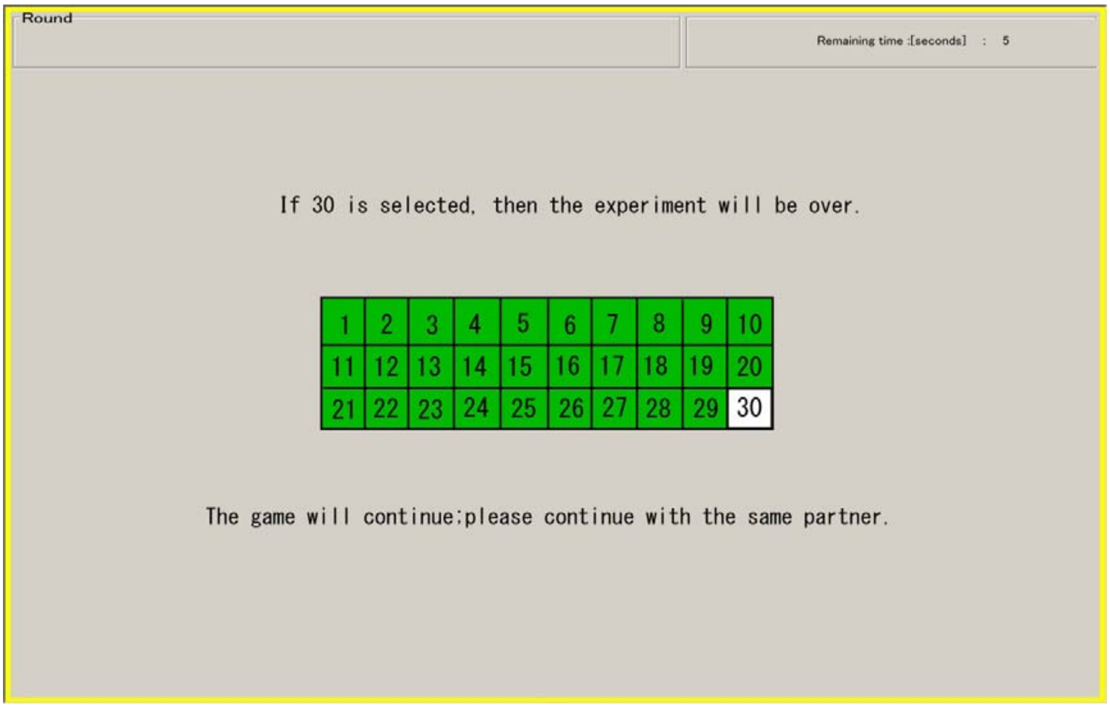

Figure A.2:

Random Termination Device

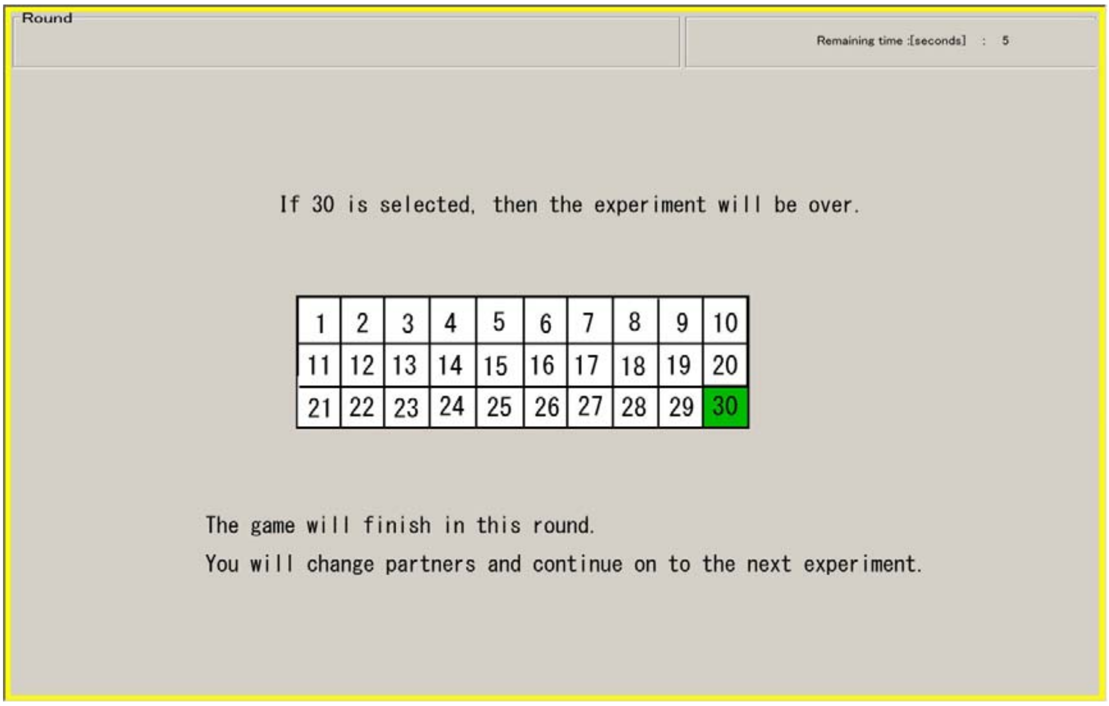

Table A.1: Mean frequencies of Cooperative Action Choices, $\rho(p)$

	p=0.9			p=0.6	
Round 1	0.781	***	(324)	0.438	(324)
Round 2-14	0.707	***	(4,212)	0.395	(4,212)
Round 15 -	0.631	***	(4,328)	0.312	(4608)
Total	0.672	***	(8,864)	0.355	(9,144)

1) Numbers of observations are in parentheses.
2) *** indicates the mean frequencies of cooperative action choices when p=0.9 are significantly higher than when p=0.6 at the 1 percent significance level by the Mann-Whitney rank-sum test.

Table A.2: Mean Frequencies of Cooperative Actions Choices on Good Signals, $r(a;p)$

	p=0.9			p=0.6	
Round 2-14	0.859	***	(2,841)	0.479	(2,072)
Round 15 -	0.844	***	(2,612)	0.397	(2,193)
Total	0.852	***	(5,453)	0.437	(4,265)

1) Numbers of observations are in parentheses.

2) *** indicates the mean frequencies of cooperative action choices when p=0.9 is significantly higher than when p=0.6 at the 1 percent level by the Mann-Whitney rank-sum test.

Table A.3: Mean Frequencies of Cooperative Action Choices on Bad Signals, $r(b; p)$

	p=0.9			p=0.6	
Round 2-14	0.392	***	(1,371)	0.314	(2,140)
Round 15 -	0.305	***	(1,716)	0.235	(2,415)
Total	0.344	***	(3,087)	0.272	(4,555)

1) Numbers of observations are in parentheses.

2) *** indicates the mean frequencies of cooperative action choices when p=0.9 is significantly higher than when p=0.6 at the 1 percent level by the Mann-Whitney rank-sum test.