

CARF Working Paper

CARF-F-452

**Behavioral Theory of Repeated Prisoner's Dilemma:
Generous Tit-For-Tat Strategy**

Hitoshi Matsushima
University of Tokyo

February 22, 2019

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

CARF Working Papers can be downloaded without charge from:
<https://www.carf.e.u-tokyo.ac.jp/research/>

Working Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Working Papers may not be reproduced or distributed without the written consent of the author.

Behavioral Theory of Repeated Prisoner's Dilemma: Generous Tit-For-Tat Strategy¹

Hitoshi Matsushima²
University of Tokyo

February 22, 2019

Abstract

This study investigates infinitely repeated games of a prisoner's dilemma with additive separability in which the monitoring technology is imperfect and private. Behavioral incentives indicate that, in this setting, a player is not only motivated by pure self-interest but also by reciprocity. Players often become naïve and select an action unconsciously. By focusing on generous tit-for-tat strategies, we characterize a Nash equilibrium with behavioral incentives, termed behavioral equilibrium, in an accuracy-contingent manner. By eliminating the gap between theory and evidence, this study argues that reciprocity plays a substantial role in motivating a player to consciously make decisions.

JEL Classification Numbers: C70, C71, C72, C73, D03.

Keywords: Repeated Prisoner's Dilemma, Additive Separability, Imperfect Private Monitoring, Generous Tit-for-Tat Strategy, Behavioral Equilibrium, Reciprocity, Naïveté.

¹ This research was financially supported by a grant-in-aid for scientific research (KAKENHI 25285059) from the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of the Japanese government, as well as by the Center of Advanced Research in Finance at the University of Tokyo. We are grateful to Yutaka Kayaba and Tomohisa Toyama for their support in drafting an earlier version of this study. All errors are mine.

² Department of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: hitoshi@e.u-tokyo.ac.jp

1. Introduction

This study examines the impact of behavioral incentives on the equilibrium outcomes in infinitely repeated games of a prisoner's dilemma with additive separability. We assume that the monitoring technology is imperfect and private. Each player cannot directly observe whether the opponent selects the cooperative or the defective action. However, each player can imperfectly monitor the opponent's action choice through private observation of a noisy signal. The signal is either "good" or "bad," and a player is more likely to observe the good signal when the opponent selects the cooperative action. In this setting, this study investigates the role of noisy signal observations on implicit collusion.

Previous works in the repeated game literature have investigated the impact of monitoring accuracy on the degree to which the history-dependence of players' strategic behaviors facilitates their implicit collusion.³ Since the monitoring technology is imperfect, it is not certain that each player will receive the good signal when the opponent makes the cooperative action choice. Therefore, monitoring imperfection inevitably interferes with the full achievement of implicit collusion.⁴

When the monitoring accuracy is satisfactory, a player can avoid the welfare loss caused by monitoring imperfection. The more accurate the monitoring technology is, the more convinced a player, who observes the bad (good) signal, is that the opponent made the defective (cooperative) action choice. Hence, the more accurate the monitoring technology, the more effectively a player can retaliate against the opponent. In other words, in the presence of high monitoring accuracy, a player can more effectively penalize the deviant even if the retaliation that he/she employs is not intensive.⁵

In contrast with this theoretical prediction, Kayaba, Matsushima, and Toyama (2019) reports experimental results indicating that subjects in laboratory experiments tend to retaliate more in the presence of accurate monitoring than in the low accuracy scenario.

³ For a survey, see Mailath and Samuelson (2006).

⁴ See Green and Porter (1984) and Abreu, Pearce, and Stacchetti (1990), for example.

⁵ The folk theorem holds even with such imperfect monitoring, thus indicating that if the discount factor is close to unity, a wide variety of allocations, including approximate efficiency, can be attained by subgame perfect equilibria (e.g., Fudenberg, Levine, and Maskin, 1994; Matsushima, 2004; Sugaya, 2019).

The expected payoff to an individual from the cooperative action choice tends to be higher than that from the defective action choice when the monitoring is accurate, while the expected payoff from the cooperative action choice tends to be lower than that from the defective action choice when the monitoring is inaccurate. These experimental findings are inconsistent with the above-mentioned theoretical predictions, thus suggesting the presence of incentives to pursue retaliation and cooperation beyond the maximization of self-interest.

This study sheds light on behavioral aspects of incentives. In particular, we argue that noisy signal observation influences the observer's psychological state, thus motivating his/her social preferences and consciousness. We assume that a player is often motivated not only by pure self-interest but also by *reciprocity*; a player feels guilty when he/she selects the defective action even though he/she observed the good signal, while a player is annoyed when he/she selects the cooperative action even if he/she observed the bad signal. In addition, we assume that a player often becomes *naïve* and selects an action unconsciously; he/she often randomly selects an action, independently of his/her pure self-interest motive and reciprocal motive.

By incorporating such reciprocity and naïveté into players' incentives, we define *behavioral equilibrium* as a natural extension of the standard Nash equilibrium notion. To simplify strategic interaction, this study will focus on generous tit-for-tat (g-TFT) strategies, which are straightforward stochastic extensions of the tit-for-tat (TFT) strategy (e.g., Molander, 1985; Nowak and Sigmund, 1992; Takahashi, 2010; Matsushima, 2013). A characterization of g-TFT behavioral equilibria in an accuracy-contingent fashion is proposed.

G-TFT is the most concise manner to describe cooperation, retaliation, and forgiveness in repeated interactions. In a g-TFT strategy, a player retaliates against the opponent by selecting the defective action more often when he/she observes the bad signal than when he/she observes the good signal. In line with this argument, the experimental studies of Kayaba, Matsushima, and Toyama (2019) indicates that among a wide variety of strategies, a significant proportion of experimental subjects adopts a g-TFT strategy even if they employ heterogeneous g-TFT.

G-TFT has a substantial advantage over (deterministic) TFT, which generally fails to be an equilibrium, while, given a high enough discount factor, g-TFT equilibria exist

irrespective of the level of monitoring accuracy. In any (accuracy-contingent) g-TFT equilibrium, the more accurate the monitoring technology is, the less intensively a player retaliates against the opponent. This view, however, contradicts the experimental results given by Kayaba, Matsushima, and Toyama (2019).

By incorporating reciprocity and naïveté into equilibrium theory, this study demonstrates a characterization result implying that, in an accuracy-contingent g-TFT behavioral equilibrium, the more accurate the monitoring technology is, the more severely each player retaliates against the opponent. This result contradicts the predictions of equilibrium theory that does not account for behavioral incentives but is consistent with the experimental evidences.

Our characterization also indicates that the more often a player behaves unconsciously (i.e., naively), the less motivated he/she is by reciprocity. Hence, reciprocity motivates a player to behave more consciously. We further show that the more accurate the monitoring technology is, the less kind a player is against the opponent: given low levels of monitoring accuracy, the less accurate the monitoring technology is, the more positively reciprocal the player is. Given a satisfactory level of monitoring accuracy, the more accurate the monitoring technology is, the more negatively reciprocal the player is. At a medium level of monitoring accuracy, a player is neither positively nor negatively reciprocal: he/she becomes the most naïve at this medium accuracy level.

This study should be regarded as the first systematic attempt in the repeated game literature to propose a behavioral theory that reconciles with experimental results. The literature of experimental repeated games has examined the relevance of theoretical predictions without behavioral incentives and the prevalence of various strategies by employing the strategy frequency estimation method (e.g., Dal Bò and Fréchette, 2011; Fudenberg, Rand, and Dreber, 2012; Aoyagi, Bhaskar, and Fréchette, 2019; Kayaba, Matsushima, and Toyama, 2019). These works commonly supported the predictions that subjects are more likely to collude as the monitoring technology is more accurate, and also indicated that subjects tend to employ heterogeneous strategies. Kayaba, Matsushima, and Toyama (2019) experimentally showed that a large proportion of subjects employ heterogeneous g-TFT strategies, and that their retaliation is severer as the monitoring technology is more accurate. Since the latter experimental observation is inconsistent with the prediction without behavioral incentives, it should be regarded as the important

research to develop a new theory that can describe this observation as equilibrium behavior. This is exactly what this study attempts to do.

Previous studies in the behavioral economics literature show that social preferences facilitate cooperation (e.g., Güth, Schmittberger, and Schwarze, 1982; Berg, Dickhaut, and McCabe, 1995; Fehr and Gächter, 2000), and preferences depend on various contexts (e.g., Rabin 1993; Charness and Rabin, 2002; Dufwenberg and Kirchsteiger, 2004; Falk and Fishbacher, 2005). This study parameterizes the relevant contexts simply by the level of monitoring accuracy. Duffy and Muñoz-García (2012) demonstrate that social preference facilitates collusion when the discount factor is insufficient. In our study, the monitoring technology is a crucial determinant of whether social preferences aid collusion. Social preferences facilitate collusion when monitoring is inaccurate, while they prevent people from colluding when monitoring is accurate.

The remainder of this study is organized as follows. Section 2 defines the repeated prisoner's dilemma with additive separability and with imperfect private monitoring. Section 3 introduces the g-TFT strategy and behavioral equilibrium. We then demonstrate our characterization result. Section 4 investigates accuracy-contingent symmetric models. Section 5 concludes.

2. Prisoner's Dilemma with Additive Separability

This study investigates an infinitely repeated *prisoners' dilemma with additive separability*, as described by Figure 1.

		player 2			
		C		D	
player 1	C	1 1	$-g_1$ $1+g_2$		
	D	$1+g_1$ $-g_2$	0 0		

Figure 1: Prisoners' Dilemma with Additive Separability

Let us call C and D the *cooperative* action and *defective* action, respectively. Because of the additive separability nature, irrespective of the opponent's action choice,

each player i generates a cost g_i by selecting the cooperative action instead of the defective action but provides the opponent $j \neq i$ a benefit equal to $1 + g_j$. We assume that $g_i > 0$ for each $i \in \{1, 2\}$, and $|g_1 - g_2| < 1$. Hence, the cooperative action profile (C, C) maximizes their total welfare, while the defective action profile (D, D) is the dominant strategy profile and is Pareto-inferior to (C, C) . Let $\delta_i \in (0, 1)$ denote the *discount factor* of player i in the associated repeated game (we allow players to have different discount factors).

We assume that monitoring is *imperfect* and *private*, as follows. Each player i cannot directly observe the action that the opponent $j \neq i$ has selected. However, he/she privately observes a noisy signal, denoted by $\omega_j \in \{c, d\}$, for the opponent j 's action choice. Let us call c and d the *good* and *bad* signals, respectively.

We define the level of *monitoring accuracy* for each player $i \in \{1, 2\}$ as a probability index $p_i \in (\frac{1}{2}, 1)$: player i observes the good signal c (the bad signal d) with probability p_j when the opponent j selects the cooperative action C (the defective action D , respectively). From $p_i > \frac{1}{2}$, the probability of receiving the good signal c for the corresponding player's action choice is higher when this player selects C than when he/she selects D . The greater p_i , the more accurately the opponent $j \neq i$ monitors player i 's action choice. To measure the level of monitoring accuracy as a single value (for simplicity), this study assumes that the probability of the good signal when the cooperative action is selected and the probability of the bad signal when the defective action is selected are equivalent. However, this assumption is irrelevant to the outcome of this study.

3. Generous Tit-For-Tat and Behavioral Equilibrium

We denote the *g-TFT strategy* for each player $i \in \{1, 2\}$ as follows:

$$s_i = (r_i(c), r_i(d)) \in [0, 1]^2.$$

In each period $t \geq 2$, player i makes the cooperative action choice C with

probability $r_i(\omega_j)$ whenever he/she observes signal $\omega_j \in \{c, d\}$ in period $t-1$. To focus on a player's incentive to make a signal-contingent action choice, we ignore incentive issues in period 1. Let us consider an arbitrary period $t \geq 2$. Each player i 's choice of C instead of D generates a cost g_i in period t , while, in period $t+1$, this choice generates a gain $1+g_i$ from the cooperative response of the opponent j with probability $p_i r_j(c) + (1-p_i)r_j(d)$ instead of probability $(1-p_i)r_j(c) + p_i r_j(d)$. Here, note that the opponent j observes the good signal with probability p_i ($1-p_i$) when player i selects the cooperative action C (the defective action D , respectively).

This study introduces a *behavioral aspect*: each player $i \in \{1, 2\}$ is motivated not only by the above-mentioned pure self-interest but also by *naïveté*, denoted by $\varepsilon_i \in [0, \frac{1}{2})$, and *reciprocity*, denoted by $(w_i(c), w_i(d)) \in R_+^2$. In every period, with probability $2\varepsilon_i$, player i becomes naïve (i.e., unconscious) and randomly selects between actions C and D . Hence, a g-TFT strategy for player i , $s_i = (r_i(c), r_i(d))$, must satisfy:

$$(1) \quad \min[r_i(c), 1-r_i(c), r_i(d), 1-r_i(d)] \geq \varepsilon_i.$$

With probability $1-2\varepsilon_i$, player i becomes conscious and makes his/her action choice in a sophisticated manner. We introduce *reciprocity* $(w_i(c), w_i(d))$ as follows. Suppose that player i observes the good signal c ; he/she feels guilty when he/she selects the defective action D despite observing the good signal. In this case, he/she can avoid a psychological cost $w_i(c) \geq 0$ by selecting the cooperative action C . The instantaneous gain from selecting action D is $g_i - w_i(c)$, while the resultant future loss is given by:

$$\begin{aligned} & \delta_i(1+g_i)\{p_i r_j(c) + (1-p_i)r_j(d)\} - \delta_i(1+g_i)\{(1-p_i)r_j(c) + p_i r_j(d)\} \\ & = \delta_i(1+g_i)(2p_i - 1)\{r_j(c) - r_j(d)\}. \end{aligned}$$

Hence, player i is willing to select the cooperative action C if:

$$g_i - w_i(c) < \delta_i(1+g_i)(2p_i - 1)\{r_j(c) - r_j(d)\},$$

while he/she is willing to select the defective action D if:

$$g_i - w_i(c) > \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\}.$$

Next, suppose that player i observes the bad signal d ; he/she is annoyed when he/she selects the cooperative action C despite observing the bad signal d . In this case, he/she can avoid a psychological cost $w_i(d) \geq 0$ by selecting the defective action D . The instantaneous gain from D is $g_i + w_i(d)$, while the resultant future loss is given by $\delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\}$. Hence, player i is willing to select the cooperative action C if:

$$g_i + w_i(d) < \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\},$$

while he/she is willing to select the defective action D if:

$$g_i + w_i(d) > \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\}.$$

From the above arguments and the additive separability nature of our component game, we define *behavioral equilibrium* as follows.

Definition 1: A g-TFT strategy profile $s \equiv (s_1, s_2)$ is said to be a *behavioral equilibrium*, or, *shortly, an equilibrium with respect to* $(g_i, \delta_i, p_i, \varepsilon_i, w_i(c), w_i(d))_{i \in \{1, 2\}}$ if, for each $i \in \{1, 2\}$:

$$(2) \quad [g_i - w_i(c) > \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\}] \Rightarrow [r_i(c) = \varepsilon_i],$$

$$(3) \quad [g_i - w_i(c) < \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\}] \Rightarrow [1 - r_i(c) = \varepsilon_i],$$

$$(4) \quad [g_i + w_i(d) > \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\}] \Rightarrow [r_i(d) = \varepsilon_i],$$

and

$$(5) \quad [g_i + w_i(d) < \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\}] \Rightarrow [1 - r_i(d) = \varepsilon_i].$$

The conditions in Definition 1 imply that if the instantaneous gain is higher (lower) than the future loss, a player selects the cooperative action (the defective action) with the minimal probability ε_i . To focus on the impact of psychological costs on incentives, let us assume that for each $i \in \{1, 2\}$:

$$\text{either } w_i(c) = 0 \text{ or } w_i(d) = 0.$$

In line with this assumption, a player i is said to be *positively (negatively) reciprocal* if

$w_i(c) > 0$ ($w_i(d) > 0$). A player i is said to be *null-reciprocal* if $w_i(c) = w_i(d) = 0$.

We define the *retaliation intensity* of player i as the difference in cooperation rate between the good and bad signal scenario:

$$r_i(c) - r_i(d).$$

The retaliation intensity measures the degree to which a player punishes the opponent when he/she observes the bad signal. Let us define

$$R_i = R_i(p_j) \equiv \frac{g_j}{\delta_j(1+g_j)(2p_j-1)}.$$

Proposition 1 indicates that R_i is the retaliation intensity of player i common to all equilibria when the opponent $j \neq i$ is null-reciprocal.

Proposition 1⁶: *Consider an arbitrary g -TFT equilibrium s and an arbitrary player $i \in \{1, 2\}$. Suppose that the opponent $j \neq i$ is null-reciprocal, such that:*

$$w_j(c) = w_j(d) = 0.$$

Then, the retaliation intensity of player i is equivalent to R_i , such that:

$$(6) \quad r_i(c) - r_i(d) = R_i(p_j).$$

Proof: Selecting C instead of D generates a cost g_i for player i in the current round, whereas, in the next round, he/she can gain $1 + g_i$ from opponent j 's response with probability $p_i r_j(c) + (1 - p_i) r_j(d)$ instead of $(1 - p_i) r_j(c) + p_i r_j(d)$. Since player i is incentivized to select both C and D , it follows from additive separability that the indifference between these action choices must be a necessary and sufficient condition for equilibrium:

$$-g_i + \delta_i(1 + g_i)\{p_i r_j(c) + (1 - p_i) r_j(d)\} = \delta_i(1 + g_i)\{(1 - p_i) r_j(c) + p_i r_j(d)\},$$

or, equivalently,

⁶ The proof of this proposition descends from the 'belief-free' nature (e.g., Ely and Välimäki, 2002; Piccione, 2002). We retain the proof in the text for the merit of self-contained treatments.

$$g_i = \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\},$$

thus implying (6).

Q.E.D.

Theorem 2 demonstrates an important characterization result and clarifies the relation between behavioral aspects and equilibrium.

Theorem 2: *Suppose that each player $i \in \{1, 2\}$ is not null-reciprocal. Consider an arbitrary g -TFT strategy profile s , where for each $i \in \{1, 2\}$, $r_i(c)$, $r_i(d)$, $1 - r_i(c)$, and $1 - r_i(d)$ are all different, and equality (6) does not hold. Then, s is an equilibrium if and only if for each $i \in N$, equality (1) and the following properties hold; if:*

$$r_j(c) - r_j(d) < R_j(p_i),$$

then:

$$(7) \quad w_i(c) = g_i - \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\} > 0,$$

$$w_i(d) = 0,$$

and

$$\varepsilon_i = r_i(d).$$

If:

$$r_j(c) - r_j(d) > R_j(p_i),$$

then:

$$w_i(c) = 0,$$

$$(8) \quad w_i(d) = \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\} - g_i > 0,$$

and

$$\varepsilon_i = 1 - r_i(c).$$

Proof: The proof of the “if” part is a direct consequence of Definition 1. The proof of the “only if” part is as follows. Suppose that:

$$r_j(c) - r_j(d) < R_j(p_i).$$

Then, the left-hand side of (4) holds; as a result:

$$\varepsilon_i = r_i(d).$$

Since $r_i(c)$ and $1 - r_i(c)$ are different from ε_i , it follows from (2) and (3) that:

$$g_i - w_i(c) = \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\},$$

thus implying (7). Since $w_i(c)$ is positive, player i is positively reciprocal, meaning that $w_i(d) = 0$.

Next, suppose that:

$$r_j(c) - r_j(d) > R_j(p_j).$$

Then, the left-hand side of (3) holds; as a result:

$$\varepsilon_i = 1 - r_i(c).$$

Since $r_i(d)$ and $1 - r_i(d)$ are different from ε_i , it follows from (4) and (5) that:

$$g_i + w_i(d) = \delta_i(1 + g_i)(2p_i - 1)\{r_j(c) - r_j(d)\},$$

thus implying (8). Since $w_i(d)$ is positive, player i is negatively reciprocal, meaning that $w_i(c) = 0$.

Q.E.D.

From Theorem 2, player i is positively reciprocal (i.e., $w_i(c) > 0$ and $w_i(d) = 0$) if the retaliation intensity is lower than R_i , while he/she is negatively reciprocal (i.e., $w_i(c) = 0$ and $w_i(d) > 0$) if the retaliation intensity is higher than R_i . The more negatively reciprocal (or the less positively reciprocal) a player is, the greater the opponent's retaliation intensity is. Moreover, from Theorem 2, we can automatically derive the unique equilibrium as follows.

Corollary 3: Consider an arbitrary behavioral equilibrium s and an arbitrary player $i \in \{1, 2\}$. If the opponent $j \neq i$ is positively reciprocal, then:

$$r_i(c) = \frac{g_j - w_j(c)}{\delta_j(1 + g_j)(2p_j - 1)} + \varepsilon_i,$$

$$r_i(d) = \varepsilon_i,$$

where the retaliation intensity is given by:

$$r_i(c) - r_i(d) = \frac{g_j - w_j(c)}{\delta_j(1 + g_j)(2p_j - 1)}.$$

If opponent j is negatively reciprocal, then:

$$r_i(c) = 1 - \varepsilon_i,$$

$$r_i(d) = 1 - \varepsilon_i - \frac{g_j + w_j(d)}{\delta_j(1 + g_j)(2p_j - 1)},$$

where the retaliation intensity is given by:

$$r_i(c) - r_i(d) = \frac{g_j + w_j(d)}{\delta_j(1 + g_j)(2p_j - 1)}.$$

4. Symmetry

This section considers a *symmetric* model in which there exists (g, δ, p) such that

$$(g_i, \delta_i, p_i) = (g, \delta, p) \text{ for each } i \in \{1, 2\}.$$

We allow players' behavioral aspects to be heterogeneous and contingent on a common level of monitoring accuracy p ; therefore, $(\varepsilon_i(p), w_i(c; p), w_i(d; p))$ instead of $(\varepsilon_i, w_i(c), w_i(d))$.

In line with Kayaba, Matsushima, and Toyama (2019), we allow players' g-TFT strategies to be heterogeneous. Moreover, an equilibrium should be contingent on the level of monitoring accuracy p ; therefore, $s_i(p) = (r_i(c; p), r_i(d; p))$ instead of $s_i = (r_i(c), r_i(d))$.

Let us set an arbitrary level $\underline{p} \in (1/2, 1)$, which we call the minimum level. This section assumes that both the accuracy-contingent behavioral aspect $(\varepsilon_i(p), w_i(c; p), w_i(d; p))$ and the accuracy-contingent equilibrium strategy $s_i(p) = (r_i(c; p), r_i(d; p))$ are *continuous* in $p \in [\underline{p}, 1]$.

4.1. Kindness and Retaliation Intensity

As a measure of the degree to which the accuracy-contingent reciprocity $(w_i(c; p), w_i(d; p))$ motivates player i to make the cooperative action choice, this section considers the following notion of *kindness*.

Definition 2: Player i is said to be *more kind at p than at \tilde{p}* if:

$$\text{either } w(c; p) > w(c; \tilde{p}) \text{ or } w(d; p) < w(d; \tilde{p}).$$

A kind player is rarely annoyed at the opponent when observing the bad signal, and he/she even praises the opponent when observing the good signal. Clearly, player i is more kind at p than at \tilde{p} if he/she is positively reciprocal at p but is negatively reciprocal at \tilde{p} .

From Proposition 1, if both players are null-reciprocal irrespective of p , then, the associated equilibrium retaliation intensity, which is common to both players, is given by:

$$R(p) \equiv \frac{g}{\delta(1+g)(2p-1)}$$

and is decreasing in p . However, in contrast with this prediction from null-reciprocity, Kayaba, Matsushima, and Toyama (2019) report experimental results showing that the retaliation intensities observed in laboratories are *increasing* in the level of monitoring accuracy. Theorem 4 indicates that the role of players' behavioral aspects is essential for explaining this experimental observation.

Theorem 4: *Suppose that $s(p)$ is an accuracy-continent equilibrium. If the retaliation intensity of player i , $r_i(c; p) - r_i(d; p)$, is increasing in $p \in [\underline{p}, 1]$, then, the higher the monitoring accuracy p is, the less kind opponent j is.*

Proof: Since $r_i(c; p) - r_i(d; p)$ is increasing and $R(p)$ is decreasing in p , it follows from Theorem 2 that there is a unique $\hat{p}_j \in [\underline{p}, 1]$ that satisfies the following properties for each $p \in [\underline{p}, 1]$; if $p < \hat{p}_j$, then:

$$r_i(c; p) - r_i(d; p) < R(p),$$

and, therefore, $w_j(c; p)$ is positive and decreasing in $p \in [\underline{p}, \hat{p}_i]$. If $p > \hat{p}_i$, then:

$$r_i(c; p) - r_i(d; p) > R(p),$$

and, therefore, $w_j(d; p)$ is positive and increasing in $p \in [\hat{p}_i, 1]$. These properties imply that the higher p is, the less kind opponent j is.

Q.E.D.

Theorem 4 implies that in the presence of high enough levels of monitoring accuracy, a player tends to be more negatively reciprocal as the accuracy of monitoring increases. This tendency makes the retaliation intensity more severe and works against the success in cooperation induced by the improvement of the monitoring technology. In the presence of low enough levels of monitoring accuracy, a player tends to be more positively reciprocal as the accuracy of monitoring decreases. This tendency makes the retaliation intensity milder and mitigates the lack of cooperation caused by the deterioration of the monitoring technology.

4.2. Naïveté and Reciprocity

Kayaba, Matsushima, and Toyama (2019) also report experimental results showing that the more likely experimental subjects are to make the cooperative action choice, the more accurate the monitoring technology is. Based on this finding, this section considers an accuracy-contingent equilibrium $s(p)$ such that for each $i \in \{1, 2\}$, $r_i(c; p)$, $r_i(d; p)$, and $r_i(c; p) - r_i(d; p)$ are all continuous and increasing in p .

Theorem 5: *Consider an arbitrary accuracy-contingent equilibrium $s(p)$. Suppose that for each $i \in \{1, 2\}$, $r_i(c; p)$, $r_i(d; p)$, and $r_i(c; p) - r_i(d; p)$ are continuous and increasing in p . Then, for each $i \in N$, $\varepsilon_i(p)$ is increasing, $w_i(c; p)$ is decreasing, and $w_i(d; p) = 0$ in $p \in [\underline{p}, \hat{p}_i]$, while $\varepsilon_i(p)$ is decreasing, $w_i(d; p)$ is increasing, and $w_i(c; p) = 0$ in $p \in [\hat{p}_i, 1]$.*

Proof: From Corollary 3, if $p < \hat{p}_i$, then:

$$\varepsilon_i(p) = r_i(d; p).$$

Since $r_i(d; p)$ is increasing in p , $\varepsilon_i(p)$ is increasing in $p \in [\underline{p}, \hat{p}_i]$. From Corollary 3, if $p > \hat{p}_i$, then:

$$\varepsilon_i(p) = 1 - r_i(c; p).$$

Since $r_i(c; p)$ is increasing in p , $\varepsilon_i(p)$ is decreasing in $p \in [\hat{p}_i, 1]$. From these observations and Theorem 2, we obtain the proof of Theorem 5.

Q.E.D.

Theorem 5 shows an important trade-off between naïveté and reciprocity; *the more likely a player is to behave naively (i.e., unconsciously), the less reciprocal he/she is.* When a player is either more negatively reciprocal or more positively reciprocal, he/she tends to be more conscious. A player's conscious decision-making is motivated not by his/her pure self-interest but also by reciprocity.

5. Concluding Remarks

This study incorporated reciprocity and naïveté into infinitely repeated prisoner's dilemma with imperfect private monitoring. The strategic behavior is described as a behavioral equilibrium that is consistent with the experimental evidences. This study is the first systematic analysis of repeated prisoner's dilemma that fills the gap between theory and evidence. We have several issues that are left unsolved as possible future research. For instance, this study focused on g-TFT strategies; however, other types of strategies such as grim-trigger, lenience, and long-term punishment, which are prominent from theoretical and empirical viewpoints, should also be investigated. This study should be extended to more general repeated games beyond repeated prisoner's dilemma. The impact of reciprocity and naïveté on implicit collusion should be investigated more directly, and more carefully, than previous works, through creating a new experimental design.

References

- Abreu, D., D. Pearce, and E. Stacchetti (1990): “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring,” *Econometrica* 58, 1041–1063.
- Aoyagi, M., V. Bhaskar, and G. Fréchette (2019): “The Impact of Monitoring in Infinitely Repeated Games: Perfect, Public, and Private,” *American Economic Journal: Microeconomics* 11(1), 1-43.
- Berg, J., J. Dickhaut, and K. McCabe (1995): “Trust, Reciprocity, and Social History,” *Games and Economic Behavior* 10, 122–142.
- Charness, G., and M. Rabin (2002): “Understanding Social Preferences with Simple Tests,” *Quarterly Journal of Economics* 117(3), 817–869.
- Duffy, J., and F. Muñoz-García (2012): “Patience or Fairness? Analyzing Social Preferences in Repeated Games,” *Games* 3(1), 56–77.
- Dal Bó, P. and G. Fréchette (2011): “The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence,” *American Economic Review* 101(1), 411–429.
- Dufwenberg, M., and G. Kirchsteiger (2004): “A Theory of Sequential Reciprocity,” *Games and Economic Behavior* 47(2), 268–298.
- Ely, J., and J. Välimäki (2002): “A Robust Folk Theorem for the Prisoner’s Dilemma,” *Journal of Economic Theory* 102, 84–105.
- Falk, A., and U. 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, E., and S. Gächter (2000): “Fairness and Retaliation: The Economics of Reciprocity,” *Journal of Economic Perspectives* 14(3), 159–181.
- Fudenberg, D., D. Levine, and E. Maskin (1994): “The Folk Theorem with Imperfect Public Information,” *Econometrica* 62, 997–1040.
- Fudenberg, D., D.G. Rand, and A. Dreber (2012): “Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World,” *American Economic Review* 102(2), 720–749.
- Green, E. and R. Porter (1984): “Noncooperative Collusion under Imperfect Price Information,” *Econometrica* 51, 87-100.
- Güth, W., R. Schmittberger, and B. Schwarze (1982): “An Experimental Analysis of

- Ultimatum Bargaining,” *Journal of Economic Behavior and Organization* 3, 367–388.
- Kayaba, Y., H. Matsushima, and T. Toyama (2019): “Accuracy and Retaliation in Repeated Games with Imperfect Private Monitoring: Experiments,” Discussion Paper CARF-F-433, University of Tokyo.
- Mailath, J. and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.
- Matsushima, H. (2004): “Repeated Games with Private Monitoring: Two Players,” *Econometrica* 72, 823–852.
- Matsushima, H. (2013): “Interlinkage and Generous Tit-For-Tat Strategy,” *Japanese Economic Review* 65, 116–121.
- Molander, P. (1985): “The Optimal Level of Generosity in a Selfish Uncertain Environment,” *Journal of Conflict Resolution* 29, 611–618.
- Nowak, M., and K. Sigmund (1992): “Tit-For-Tat in Heterogeneous Populations,” *Nature* 355, 250–253.
- Piccione, M. (2002): “The Repeated Prisoners’ Dilemma with Imperfect Private Monitoring,” *Journal of Economic Theory* 102, 70–83.
- Rabin, M. (1993): “Incorporating Fairness into Game Theory and Economics,” *American Economic Review* 83(5), 1281–1302.
- Sugaya, T. (2019): *The Folk Theorem in Repeated Games with Private Monitoring*, mimeo.
<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx0YWwt1b3N1Z2F5YXxneDozNjE1YTM2OGQwZTM3M2Ex>
- Takahashi, S. (2010): “Community Enforcement when Players Observe Partners’ Past Play,” *Journal of Economic Theory* 145, 42–64.