数理解析研究所講究録
1488 巻 2006 年 156-176

Efficiency and Diversity in Voluntarily Repeated Prisoner's Dilemma*

Takako Fujiwara-Greve
Department of Economics,
Keio University
2-15-45 Mita, Minatoku, Tokyo 108-8345 Japan
Email: takakofg@econ.keio.ac.jp
and
Masahiro Okuno-Fujiwara
Department of Economics,
University of Tokyo
7-3-1 Hongo, Bunkyoku, Tokyo 113-0033 Japan
Email: fujiwara@e.u-tokyo.ac.jp

Abstract

We develop a general framework to analyze endogenous relationships. To consider relationships in the modern society, neither one-shot games nor repeated games are appropriate models because the formation and dissolution of a relationship is not an option. We formulate voluntarily separable repeated games, in which players are randomly matched to play a component game as well as to choose whether to play the game again with the same partner. When the component game is a prisoner's dilemma, neutrally stable distribution (NSD) requires some *trust-building periods* to defect at the beginning of a partnership. We find that *bimorphic* NSDs with voluntary break-ups include strategies with shorter trust-building periods than any *monomorphic* NSD with no voluntary separation, and hence the average payoff of bimorphic NSD is higher.

Keywords: voluntary separation, prisoner's dilemma, evolution, trust.
JEL classification: C73

1 Introduction

We develop a general framework to analyze endogenous relationships. To consider relationships in the modern society, neither one-shot games nor repeated games are appropriate models because the formation and dissolution of a relationship is not an option.

*Special thanks go to Nobue Suzuki, for her constant help. We also thank useful comments from Akihiko Matsui, Chiaki Hara, and seminar participants at KIER, Kyoto University, Keio University, University of Tokyo, Mathematical Sciences Conference (Kyoto University, November 2005), and Symposium on Market Quality (Keio Univeristy, Tokyo, December 2005). All errors are ours. This work was supported by Grant-in-Aid for Scientific Research on Priority Areas, No.12124204.

We formulate voluntarily separable repeated games in a large society of homogeneous players. Players are randomly matched to play a stage game, and. after each round of play, they can choose whether to continue playing the game with the same partner or not. Each direct interaction (a partnership) is voluntarily separable, and, moreover, there is no information flow to other partnerships.

We focus on two-person prisoner's dilemma as the component game, since it highlights the merit of mutual cooperation as well as a strong incentive to defect and escape to avoid retaliation. There are many real-world situations which fit this model. Borrowers can move from a city to another after defaulting. Workers can shirk and then quit the job. Still, we often observe cooperative modes of behavior in such situations. We provide an evolutionary foundation to cooperative behaviors. We extend Neutrally Stable Distribution (NSD) concept to fit for our model, under which no other strategy earns strictly higher payoff than the incumbents do.

Known disciplining strategies such as trigger strategies (Fudenberg and Maskin, 1986) and contagion of defection (Kandori, 1992, and Ellison, 1994) do not sustain cooperation in our model. There are two reasons. First, personalized punishment is impossible due to the ability to end the partnership unilaterally and the lack of information flow to the future partners. Second. the large society and random death make it impossible to spread defection in the society to eventually reach the original deviator. Our model describes a large, anonymous, and member-changing society, which needs a different type of discipline from those of a society of directly interacting long-run players.

Some literature exists on voluntarily separable repeated games for generalized prisoner's dilemma (Datta, 1996, Kranton 1996a, and Ghosh and Ray, 1997). They focused on symmetric strategy distributions in which all (rational) players play the same strategy and showed that a gradual-cooperation strategy sustains eventual cooperation. By contrast, our framework is valid for any component game, and we consider both symmetric (called *monomorphic*) strategy distributions, in which no voluntary separation occurs, and fundamentally asymmetric strategy distributions, in which voluntary separation occurs on the equilibrium play path.

We first show that Defect must be played initially to sustain future cooperation. We then identify a relationship between the death rate (discount factor) and the sufficient number of initial defection (called *trust-building periods*) of both monomorphic NSDs and bimorphic NSDs. We found that bimorphic NSDs include strategies with shorter trust-building periods than monomorphic NSDs, thanks to double disciplining by not only trust building but also possible exploitation by a strategy with longer trust-building periods. Hence bimorphic NSDs are more efficient than the most efficient monomorphic NSD.

2 Model and Stability Concepts

2.1 Model

Consider a society with a continuum of players, each of whom may die in each period $1, 2, \ldots$ with probability $0 < (1 - \delta) < 1$. When they die, they are replaced by newly born players, keeping the total population constant. A newly born player enters into the *matching pool* where players are randomly paired to play a *Voluntarily Separable Prisoner's Dilemma (VSPD)* as follows.[1]

In each period, players play the following *Extended Prisoners' Dilemma* (EPD). First, they play ordinary one-shot prisoners' dilemma, whose actions are denoted as Cooperate and Defect. After observing the play action profile of the period by the two players, they choose simultaneously whether or not they want to keep the match into the next period (action k) or bring it to an end (action e). Unless both choose k, the match is dissolved and players will have to start the next period in the matching pool. In addition, even if they both choose k, partner may die with probability $1 - \delta$ which forces the player to go back to the matching pool next period. If both choose k and survive to the next period, then the match continues, and the matched players play EPD again.

Assume that there is limited information available to play EPD. In each period, players know the VSPD history of their current match but have no knowledge about the history of other matches in the society.

[1] Although we focus on Prisoner's Dilemma as the component game, the framework can be applied to any component game.

TABLE I

PAYOFF OF PD

P1 \ P2	C	D
C	c, c	ℓ, g
D	g, ℓ	d, d

In each match, a profile of play actions determines the players' instantaneous payoffs for each period while they are matched. We denote the payoffs associated with each play action profile as: $u(C,C) = c$, $u(C,D) = \ell$, $u(D,C) = g$, $u(D,D) = d$ with the ordering $g > c > d > \ell$ and $2c \geq g + \ell$. (See Table I.)

Because we assume that the innate discount rate is zero except for the possibility of death, each player finds the relevant discount factor to be $\delta \in (0,1)$. With this, life-long payoff for each player is well-defined given his own strategy (for VSPD) and the strategy distribution in the matching pool population over time.

Let $t = 1, 2, \ldots$ indicate the periods in a match, not the calendar time in the game. Under the limited information assumption, without loss of generality we can focus on strategies that only depend on t and the private history of actions in the Prisoner's Dilemma within a match.[2] The (infinite) set of pure strategies of VSPD is denoted as $\boldsymbol{S}$ and the set of all strategy distributions in the population is denoted as $\mathcal{P}(\boldsymbol{S})$. For simplicity we assume that each player uses a pure strategy.

We investigate stability of stationary strategy distributions in the matching pool. Although the strategy distribution in the matching pool may be different from the distribution in the entire society, if the former is stationary, the distribution of various states of matches (strategy pair and the "age" of the partnership) is also stationary, thanks to the stationary death process. Hence stability of stationary strategy distributions in the matching pool implies stability of "social states". By looking at the strategy distributions in the matching pool, we can directly compute life-time payoffs of players easily.

When a strategy $s \in \boldsymbol{S}$ is matched with another strategy $s' \in \boldsymbol{S}$, the *expected length* of the match is denoted as $L(s, s')$ and is computed as follows. Notice that even if s and

[2]The continuation decision is observable, but strategies cannot vary depending on combinations of $\{k, e\}$ since only (k, k) will lead to the future choice of actions.

s' intend to maintain the match, it will only continue with probability δ^2, which is the probability that both survive to the next period. Suppose that if no death occurs while they form the partnership, s and s' will end the partnership at the end of $T(s,s')$-th period of the match. Then

$$L(s,s') := 1 + \delta^2 + \delta^4 + \cdots + \delta^{2\{T(s,s')-1\}} = \frac{1-\delta^{2T(s,s')}}{1-\delta^2}.$$

The *expected total discounted value of the payoff stream of s within the match with s'* is denoted as $V^I(s,s')$. The *average per period payoff* that s expects to receive within the match with s' is denoted as $v^I(s,s')$. Clearly,

$$v^I(s,s') := \frac{V^I(s,s')}{L(s,s')}, \text{ or } V^I(s,s') = L(s,s')v^I(s,s').$$

The expected life-time payoff of a strategy $s \in \boldsymbol{S}$ when the matching pool has a stationary distribution $p \in \mathcal{P}(\boldsymbol{S})$ is denoted as $V(s;p)$. A straightforward way to compute $V(s;p)$ is to set up a recursive equation. If p has a finite support, then we can write

$$\begin{aligned} V(s;p) &= \sum_{s' \in supp(p)} p(s')\Big[V^I(s,s') \\ &\quad + [\delta(1-\delta)\{1+\delta^2+\cdots+\delta^{2\{T(s,s')-2\}}\} + \delta^{2\{T(s,s')-1\}}\delta]V(s;p)\Big], \end{aligned}$$

where $supp(p)$ is the support of the distribution p, $T(s,s')$ is the date at the end of which s and s' end the match, the sum $\delta(1-\delta)\{1+\delta^2+\cdots+\delta^{2\{T(s,s')-2\}}\}$ is the probability that s loses the partner s' before $T(s,s')$, and $\delta^{2\{T(s,s')-1\}}\delta$ is the probability that the match continued until $T(s,s')$ and s survives at the end of $T(s,s')$ and goes back to the matching pool.

Let $L(s;p) := \sum_{s' \in supp(p)} p(s')L(s,s')$. By computation,

$$\begin{aligned} V(s;p) &= \sum_{s' \in supp(p)} p(s')\Big[V^I(s,s') + \{1-(1-\delta)L(s,s')\}V(s;p)\Big] \\ &= \sum_{s' \in supp(p)} p(s')V^I(s,s') + \{1 - \frac{L(s;p)}{L}\}V(s;p). \end{aligned}$$

Hence the average payoff can be decomposed[3] as a convex combination of "in-match"

[3]However, this means that, in general, $v(s;p) \neq \sum_{s'} p(s')v^I(s,s')$. That is, v is not linear in the second component. This is due to the recursive structure of the V function.

average payoff:

$$v(s;p) = \frac{V(s;p)}{L} = \sum_{s' \in supp(p)} p(s') \frac{L(s,s')}{L(s;p)} v^I(s,s'), \tag{1}$$

where the ratio $L(s,s')/L(s;p)$ is the relative length of periods that s expects to spend in a match with s'. In particular, if p is a strategy distribution consisting of a single strategy s', then

$$v(s;p) = v^I(s,s').$$

2.2 Nash Equilibrium

DEFINITION. Given a stationary strategy distribution in the matching pool $p \in \mathcal{P}(\boldsymbol{S})$, $s \in \boldsymbol{S}$ is a *best reply against* p if for all $s' \in \boldsymbol{S}$,

$$v(s;p) \geq v(s';p),$$

and is denoted as $s \in BR(p)$.

DEFINITION. A stationary strategy distribution in the matching pool $p \in \mathcal{P}(\boldsymbol{S})$ is a *Nash equilibrium* if, for all $s \in \text{supp}(p)$, $s \in BR(p)$.

LEMMA 1. *For any pure strategy $s \in \boldsymbol{S}$ that starts with C in $t = 1$, let p_s be the strategy distribution consisting only of s. Then p_s is not a Nash equilibrium.*

PROOF: Consider a myopic strategy $\tilde{d}$ as follows.

$t = 1$: Play D and e (end the partnership) for any observation.

$t \geq 2$: Since this is off-path, any action can be specified.

Clearly, $\tilde{d}$-strategy earns g as the average payoff under p_s, which is the maximal possible payoff. I.e., $\tilde{d} \in BR(p_s)$ and $s \notin BR(p_s)$. Q.E.D.

Therefore, trigger strategy used in the ordinary folk theorem of repeated prisoner's dilemma cannot constitute even a Nash equilibrium. There needs to be at least one period of (D,D) in any symmetric equilibrium.

By contrast, $p_{\tilde{d}}$ consisting only of $\tilde{d}$-strategy is a Nash equilibrium. Against $\tilde{d}$, any strategy must play one-shot PD. Hence, any strategy that starts with C in $t = 1$ earns strictly lower average payoff than that of $\tilde{d}$, and any strategy that starts with D in $t = 1$ earns the same average payoff as that of $\tilde{d}$.

2.3 Neutral Stability

Recall that in an ordinary 2-person symmetric normal-form game $G = (S, u)$, a (mixed) strategy $p \in \mathcal{P}(S)$ is a Neutrally Stable Strategy if for any $q \in \mathcal{P}(S)$, there exists $0 < \bar{\epsilon}_q < 1$ such that for any $\epsilon \in (0, \bar{\epsilon}_q)$, $Eu(p, (1-\epsilon)p + \epsilon q) \geq Eu(q, (1-\epsilon)p + \epsilon q)$. (Maynard Smith, 1982.)

An extension of this concept to our extensive form game is to require a strategy distribution not to be invaded by a small fraction of a mutant strategy who enters the matching pool in a stationary manner.

DEFINITION. Given $\epsilon > 0$ and a (stationary) strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$ in the matching pool, a strategy $s' \in \boldsymbol{S}$ *invades* p for ϵ if for any $s \in supp(p)$,

$$v(s'; (1-\epsilon)p + \epsilon p_{s'}) \geq v(s; (1-\epsilon)p + \epsilon p_{s'}), \tag{2}$$

and for some $s \in supp(p)$,

$$v(s'; (1-\epsilon)p + \epsilon p_{s'}) > v(s; (1-\epsilon)p + \epsilon p_{s'}), \tag{3}$$

where $p_{s'}$ is the strategy distribution consisting only of s'.

A weaker notion of invasion that requires weak inequality only (which is used in the notion of Evolutionary Stable Strategy) is too weak in our extensive-form model since any strategy that is different in the off-path actions from the incumbent strategies can invade under the weak inequality condition.

DEFINITION. A (stationary) strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$ in the matching pool is a *Neutrally Stable Distribution* (NSD) if, for any $s' \in \boldsymbol{S}$, there exists $\bar{\epsilon} \in (0, 1)$ such that s' cannot invade p for any $\epsilon \in (0, \bar{\epsilon})$.

If a symmetric strategy distribution consisting of a single pure strategy s is a neutrally stable distribution, then s is called a Neutrally Stable Strategy (NSS). The condition for s

to be a NSS reduces to: for any $s' \in \boldsymbol{S}$, there exists $\bar{\epsilon} \in (0,1)$ such that, for any $\epsilon \in (0, \bar{\epsilon})$,

$$v(s; (1-\epsilon)p_s + \epsilon p_{s'}) \geq v(s'; (1-\epsilon)p_s + \epsilon p_{s'}).$$

It can be easily seen that any NSD is a Nash equilibrium.

Similar to the "static" notion of evolutionary stability, this definition is based on the assumption that mutation takes place rarely so that only single mutation occurs within the time span in which stationary strategy distribution is formed. However, unlike the ordinary notion of neutral stability (or ESS) of one-shot games, we need to assume the expected length of the life-time of a mutant strategy in order to calculate the average payoff. We adopted a strong requirement that the incumbents are not worse-off than mutants even if mutants stay stationarily in the population, let alone if they die out. While we do not insist that the above definition is the best among we can imagine, it is tractable and justifiable.

We now show that $\tilde{d}$-strategy is not NSS, even though it constitutes a symmetric Nash equilibrium.

LEMMA 2. *Myopic $\tilde{d}$-strategy is not an NSS.*

PROOF: Consider the following c_1-strategy.

$t = 1$: Play D and keep the partnership if and only if (D, D) is observed in the current period.

$t \geq 2$: Play C and keep the partnership if and only if (C, C) is observed in the current period.

For any $\epsilon \in (0,1)$, let $p := (1-\epsilon)p_{\tilde{d}} + \epsilon p_1$. From (1),

$$\begin{aligned} v(\tilde{d}; p) &= d; \\ v(c_1; p) &= (1-\epsilon)\frac{L(c_1, \tilde{d})}{L(c_1; p)} v^I(c_1, \tilde{d}) + \epsilon \frac{L(c_1, c_1)}{L(c_1; p)} v^I(c_1, c_1) > d, \end{aligned}$$

since $v^I(c_1, \tilde{d}) = d$, and $v^I(c_1, c_1) = (1-\delta^2)d + \delta^2 c > d$. Q.E.D.

3 Monomorphic Equilibria

We will analyze equilibria of a certain form called *trust-building strategies*. Our purpose of this paper is not to provide a folk theorem but to clarify how repeated cooperation can be attained by boundedly rational players in anonymous society who do not play carefully constructed punishment strategies. Needless to say, Nash equilibrium and NSD are proved by checking all other strategies in $\boldsymbol{S}$ (not just among trust-building strategies).

DEFINITION. For any $T = 1, 2, 3, \ldots$, let a *trust-building strategy* with T periods of trust-building (written as c_T-strategy hereafter) be a strategy satisfying

$t = 1, \ldots, T$: Play D and keep the partnership if and only if (D, D) is observed in the current period.

$t \geq T+1$: Play C and keep the partnership if and only if (C, C) is observed in the current period.

The first T periods of c_T-strategy are called *trust-building periods* and the periods afterwards are called the *cooperation periods*. This class of strategies are of particular interest, since if matched players use the same c_T-strategy, the cooperation periods give the most efficient symmetric outcome as long as they live. However, in order to sustain the perpetual cooperation, we need at least one period of (D, D) due to Lemma 1. We are interested in the shortest trust-building periods to sustain such a cooperative long-term relationship.

Let p_T be the strategy distribution consisting only of c_T-strategy. The average payoff of c_T-strategy when p_T is the stationary strategy distribution in the matching pool is computed as follows. A match of c_T against c_T continues as long as they both live and the payoff sequence is d for the first T periods and c thereafter:

$$\begin{aligned} L(c_T, c_T) &= 1 + \delta^2 + \cdots = \frac{1}{1 - \delta^2}, \\ V^I(c_T, c_T) &= \{1 + \delta^2 + \cdots + \delta^{2(T-1)}\}d + (\delta^{2T} + \cdots)c. \end{aligned}$$

Since $v(c_T; p_T) = v^I(c_T, c_T) = \frac{V^I(c_T, c_T)}{L(c_T, c_T)}$, the average payoff is

$$v(c_T; p_T) = (1 - \delta^{2T})d + \delta^{2T}c. \tag{4}$$

By the logic of dynamic programming, it is necessary and sufficient for a strategy to be optimal that it cannot be improved by one-step deviations. Although the literal one-step deviation is infeasible in our model (since a player cannot change strategies across matches), it is easy to see that if a strategy is unimprovable by (infeasible) one-step deviations, then it is unimprovable by any strategy within $\boldsymbol{S}$. Therefore we find a condition that strategies which differ from c_T in one-step (in particular during the cooperation periods) do not give a higher average payoff than c_T-strategy when the stationary strategy distribution in the matching pool is p_T.

Consider a strategy that plays D at some point during the cooperation periods. It receives g as the immediate payoff but returns to the matching pool immediately if the player does not die. The payoff is $g + \delta V(c_T; p_T)$. On the other hand, the expected continuation payoff of c_T-strategy during the cooperation periods is $L(c_T, c_T)c + \delta(1-\delta)(1+\delta^2+\cdots)V(c_T; p_T) = L(c_T, c_T)c + \delta(1-\delta)L(c_T, c_T)V(c_T; p_T)$. Thus, one-step deviant strategy does not earn a higher payoff than c_T-strategy if and only if

$$\begin{aligned} & g + \delta V(c_T; p_T) \leq L(c_T, c_T)c + \delta(1-\delta)L(c_T, c_T)V(c_T; p_T), \\ \iff \quad & v(c_T; p_T) = (1-\delta)V(c_T; p_T) \leq \frac{1}{\delta^2}[c - (1-\delta^2)g] =: v^{BR}, \end{aligned} \tag{5}$$

which we call the *Best Reply Condition.* Since v^{BR} is independent of the length T of trust-building periods and $v(c_T; p_T)$ decreases as T increases, (5) implies a lower bound to T.

It is straightforward to see that the Best Reply Condition (5) is the only condition that is required for p_T to be a Nash equilibrium. In the explicit expression of the parameters, the Best Reply Condition is

$$\frac{g-c}{c-d} \leq \frac{\delta^2(1-\delta^{2T})}{1-\delta^2}.$$

Given[4] T, define $\underline{\delta}(T)$ as the solution to

$$\frac{g-c}{c-d} = \frac{\delta^2(1-\delta^{2T})}{1-\delta^2}.$$

Then the Best Reply Condition (5) is satisfied if and only if $\delta \geq \underline{\delta}(T)$. It is easy to see

[4] We have implicitly fixed a G.

that

$$\underline{\delta}(1) = \sqrt{\frac{g-c}{c-d}} > \cdots > \underline{\delta}(\infty) = \sqrt{\frac{g-c}{g-d}}.$$

Although $\underline{\delta}(1)$ may exceed 1, $\underline{\delta}(\infty) < 1$. Hence for any $\delta > \underline{\delta}(\infty)$, there exists the minimum length of trust building periods that warrants (5):

$$\underline{\tau}(\delta) := argmin_{\tau \in \mathbb{R}_{++}} \{\underline{\delta}(\tau) \mid \delta \geq \underline{\delta}(\tau)\}.$$

It is easy to see that $\underline{\tau}$ is a decreasing function of δ.

PROPOSITION 1. *For any $\delta \in (\underline{\delta}(\infty), 1)$, the monomorphic strategy distribution p_T consisting only of c_T-strategy is a Nash equilibrium if and only if $T \geq \underline{\tau}(\delta)$.*

PROOF: No strategy which differ on the play path from c_T-strategy is better off if and only if T is sufficiently long so that $T \geq \underline{\tau}(\delta)$. Strategies that differ from c_T-strategy off the play path do not give a higher payoff. Q.E.D.

Note that the lower bound to the discount factor (as δ^2) that sustains the trigger-strategy equilibrium of the ordinary repeated prisoner's dilemma is $\sqrt{\frac{g-c}{g-d}} = \underline{\delta}(\infty)$. This means that cooperation in VSPD requires more patience.

Next we investigate when a Nash equilibrium p_T is neutrally stable. In general, in order to check whether a Nash equilibrium distribution is a NSD, we only need to consider mutants that are best replies to the Nash equilibrium distribution. There are only two kinds of strategies that may become alternative best replies to p_T. The obvious ones are those that differ from c_T-strategy off the play path. These will give the same payoff as c_T-strategy and therefore cannot invade p_T. The other kind is the strategies that play D at some point in the cooperation periods. When $T > \underline{\tau}(\delta)$, however, such strategies are not alternative best reply. Therefore c_T-strategy is NSS for this case.

It remains to consider the case when $\underline{\tau}(\delta)$ is an integer so that $p_{\underline{\tau}}$ is a Nash equilibrium ($\underline{\tau}$ is an abbreviation of $\underline{\tau}(\delta)$ hereafter). For this case, we have alternative best replies to $p_{\underline{\tau}}$, among which $c_{\underline{\tau}+1}$ earns the highest payoff when meeting itself. It suffices to check if $c_{\underline{\tau}+1}$-strategy cannot invade $p_{\underline{\tau}}$.

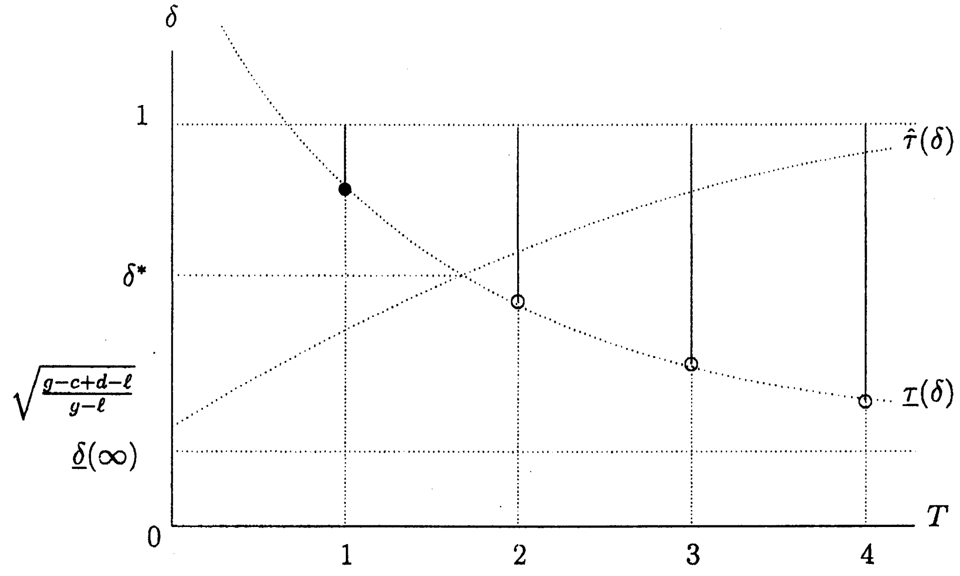

FIGURE 1. – Parametric summary of monomorphic NSS

Let us define $\hat{\tau}(\delta)$ implicitly as the solution to

$$[1 - \delta^{2(T+1)}](g - \ell) = c - d,$$

then $c_{\underline{\tau}+1}$-strategy cannot invade $p_{\underline{\tau}}$ if and only if $\underline{\tau}(\delta) < \hat{\tau}(\delta)$.

Figure 1 shows that there exists a unique $\delta^* \in (\underline{\delta}(\infty), 1)$ that satisfies

$$\delta \gtreqless \delta^* \iff \hat{\tau}(\delta) \gtreqless \underline{\tau}(\delta).$$

PROPOSITION 2. *(a) For any δ such that $\delta^* < \delta < 1$, c_T-strategy is NSS if and only if $T \geq \underline{\tau}(\delta)$.*

(b) For any δ such that $\underline{\delta}(\infty) < \delta \leq \delta^*$, c_T-strategy is NSS if and only if $T > \underline{\tau}(\delta)$.

4 Bimorphic Equilibria

The literature on voluntarily separable repeated games has concentrated on monomorphic equilibria so that no voluntary break-up occurs, except for sorting out inherent defectors under incomplete information case. We now investigate equilibria consisting of

c_T-strategies with different length of trust-building periods, hence voluntary break-ups occur on the play path. Recall that our model is of complete information and with homogeneous players. Therefore this section can be interpreted as an evolutionary foundation to the incomplete information models of diverse types of behaviors.

4.1 Existence

We investigate the shortest T for a two-strategy distribution (called *bimorphic* distribution) of $p_T^{T+1}(\alpha) = \alpha p_T + (1-\alpha)p_{T+1}$ to be a NSD for some $\alpha \in (0,1)$. For a bimorphic distribution to be a NSD, all strategies in the support must earn the same average payoff for some $\alpha \in (0,1)$. Moreover, if α increases, c_T-strategy should be worse than c_{T+1}-strategy and vice versa. Then the strategy distribution cannot be invaded by strategies that have the same play path as c_T or c_{T+1}-strategy.

Payoff Equalization: there exists $\alpha_T^{T+1} \in (0,1)$ such that

$$\alpha \gtreqless \alpha_T^{T+1} \iff v(c_{T+1}; p_T^{T+1}(\alpha)) \gtreqless v(c_T; p_T^{T+1}(\alpha)). \tag{6}$$

To derive the Best Reply Condition, note that there are two kinds of one-step deviations under a bimorphic distribution. First, a strategy can play D and keep the partnership until the partner ends the match. This strategy earns the same average payoff as c_{T+k}-strategy with $k \geq 2$. Second, a strategy can imitate c_T-strategy and play D for T periods and C at $T+1$-th period. Regardless of the partner's action, it can keep the partnership and play D in $T+2$ to earn g. (Later deviation does not earn a higher payoff due to discounting.)

Both kinds of one-step deviation do not earn higher average payoff than the incumbent c_T and c_{T+1}-strategies if and only if a similar condition to (5) as follows holds.

$$v(c_T; p_T^{T+1}(\alpha_T^{T+1})) \leq v^{BR}. \tag{7}$$

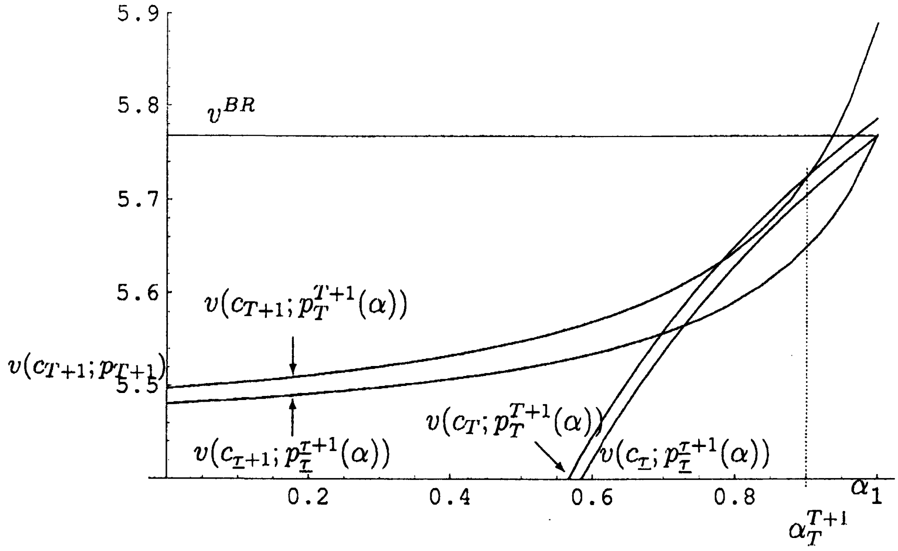

FIGURE 2. – The existence of a bimorphic NSD as T is slightly below $\underline{\tau}(\delta)$.
(Parameter values: $g = 10, c = 6.1, d = 2.1, \ell = 2, \delta = 0.96, T = 1, \underline{\tau}(\delta) \approx 1.06$.)

As before, the boundary case of $v(c_T; p_T^{T+1}(\alpha_T^{T+1})) = v^{BR}$ may not warrant a NSD but the interior case is sufficient.

Let us describe the intuition of the existence of a bimorphic NSD using Figure 2. Clearly, there is no bimorphic NSD with the support $\{c_{\underline{\tau}}, c_{\underline{\tau}+1}\}$ (where $\underline{\tau} = \underline{\tau}(\delta)$). For T *slightly* below $\underline{\tau}$, the average value functions $v(c_T; p_T^{T+1}(\alpha))$ and $v(c_{T+1}; p_T^{T+1}(\alpha))$ intersect at $\alpha < 1$ and the value at the intersection is below v^{BR}. The latter holds when the slope of $v(c_T; p_T^{T+1}(1))$ is smaller than the slope of $v(c_{T+1}; p_T^{T+1}(1))$, that is, when $T < \hat{\tau}(\delta)$.

PROPOSITION 3. *For any $\delta \geq \delta^*$, there exists $\tau^*(\delta)$ such that $\tau^*(\delta) \leq \underline{\tau}(\delta)$, and, for any $\delta > \delta^*$ and any T such that $\tau^*(\delta) < T < \underline{\tau}(\delta)$, there exists a bimoprhic NSD of the form $p_T^{T+1}(\alpha_T^{T+1}(\delta))$, where $\alpha_T^{T+1}(\delta) \in (0,1)$.*

PROOF: See Appendix.

Therefore, cooperation and exploitation can co-exist. The minimal trust-building periods $\tau^*(\delta)$ warrants that the payoff-equalizing $\alpha_T^{T+1}(\delta)$ exists. Then one can prove

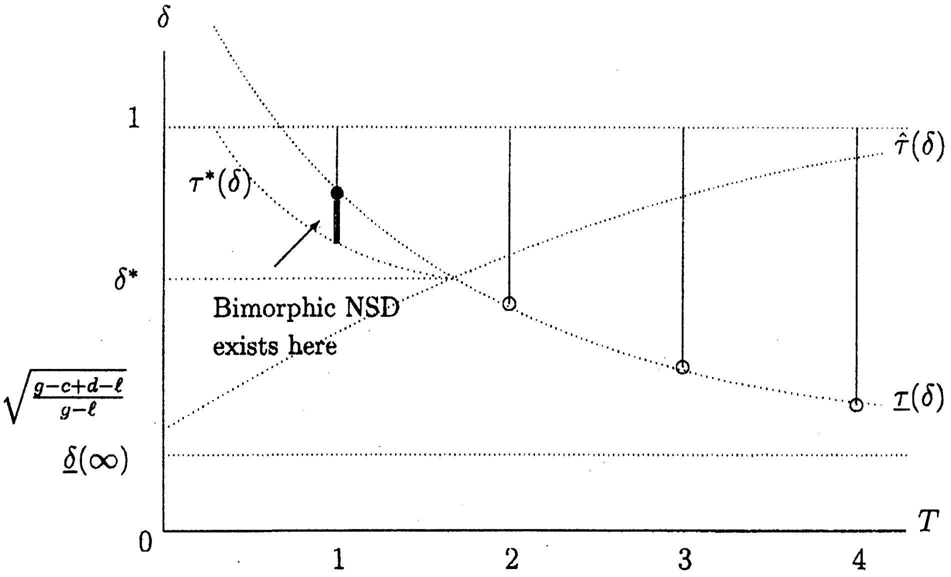

FIGURE 3. – Bimorphic NSD.

that the Best Reply Condition is satisfied for that $\alpha_T^{T+1}(\delta)$. Unlike monomorphic NSDs, however, we need δ to be sufficiently large, i.e., $\delta > \delta^*$. To warrant an integer T, we need to restrict G so that $(\tau^*(\delta), \underline{\tau}(\delta))$ contains an integer. Figure 2 is a numerical example of such G.

4.2 Higher Efficiency of Bimorphic NSD

For a given $\delta > \delta^*$, the shortest trust-building periods in the support of a bimorphic NSD is at least one period less than any of monomorphic NSS. Let the shortest trust-building periods of NSS be $T+1$ and consider a bimorphic NSD with the support $\{c_T, c_{T+1}\}$. The average payoff of c_{T+1} strategy as a NSS is

$$v(c_{T+1}; p_{T+1}) = v(c_{T+1}; p_T^{T+1}(0)).$$

Since $v(c_{T+1}; p_T^{T+1}(\alpha))$ is an increasing function of α,

$$v(c_{T+1}; p_T^{T+1}(0)) < v(c_{T+1}; p_T^{T+1}(\alpha_T^{T+1})),$$

where the latter is the average payoff of c_T and c_{T+1}-strategy under the bimorphic NSD. (See Figure 2.) Hence bimorphic NSDs, if they exist, are more efficient than any monomorphic NSS.

5 Concluding Remarks

Several papers (Datta, 1996, Kranton, 1996a, Ghosh and Ray, 1996, and Carmichael and MacLeod, 1997) have previously analyzed the voluntarily separable games, though not as fully as this paper does. These literature has pointed out two factors that facilitate cooperation under the VSPD type games.

First, they identify our symmetric trust-buidling NSD, i.e., "gradual cooperation" or "starting small" is the mechanism for sanction against defection because it makes the initial value of a new partnership small.

Our paper has more primitive structure than the papers cited above; the game is of complete information, the component game is an ordinary PD game with two actions, and there is no gift exchange stage prior to the partnership. In exchange, we develop various new concepts and a much richer set of analytical tools that enable us to investigate VSPD more fully. Furthermore, we consider evolution of behaviors within a society as a whole, rather than restricting attention to behaviors within a single partneship given (symmetric) strategy distribution in a society. As byproducts, we are able to provide fuller characterizations of symmetric trust-building strategy NSD, such as indentifying the condition (in terms of death rate and payoff values of stage game) for the existence of NSD with a particular length of trust-building periods, etc.

Second, "heterogeneity" may help cooperation. With incomplete information model, Rob and Yang (2005), independently written from ours, shows that repeated cooperation from the outset of a partnership can be an equilibrium among heterogeneous players. In their model, there are three types of players; bad type who always plays D, good type who always plays C, and rationanl type who tries to maximize their payoff. Existence of bad type players makes it valuable to (1) keep and cooperate with either good or rational type partners, and (2) to find out bad type partners as soon as possible. Thus, a rational player should cooperate from the beginning to be distinguished from the bad-type.

Our result is much starker than Rob and Yang. Our model does not rely on heterogeous "type" and incomplete information. Instead, bad (longer trust-buidling) strategy emerges endogenously as a bimorphic NSD.

Appendix: Proof of Proposition 3

We prove some useful lemmas first. For any $T, T' \in \mathbb{N}$, define

$$\Gamma(c_T, c_{T'}) := L(c_T, c_{T'})\{v^I(c_T, c_{T'}) - v^{BR}\}.$$

Then the following lemma is immediate.

LEMMA 3. *For any $T, T' \in \mathbb{N}$, if $T, T' \geq 1$, then:*

$$\Gamma(c_T, c_{T'}) = d - v^{BR} + \delta^2 \Gamma(c_{T-1}, c_{T'-1}). \tag{8}$$

Proof of Lemma 3: By definitions of Γ, L and V^I:

$$\begin{aligned}\Gamma(c_T, c_{T'}) &= L(c_T, c_{T'})v^I(c_T, c_{T'}) - L(c_T, c_{T'})v^{BR} \\ &= V^I(c_T, c_{T'}) - L(c_T, c_{T'})v^{BR} \\ &= d + \delta^2 V^I(c_{T-1}, c_{T'-1}) - \{1 + \delta^2 L(c_{T-1}, c_{T'-1})\}v^{BR} \\ &= d - v^{BR} + \delta^2 \Gamma(c_{T-1}, c_{T'-1}). \quad \square\end{aligned}$$

LEMMA 4. *For any $T \in \mathbb{N}$ and for any $v \in \mathbb{R}$:*

$$L(c_{T+1}, c_T)\{v^I(c_{T+1}, c_T) - v\} \gtreqless L(c_T, c_T)\{v^I(c_T, c_T) - v\} \iff v \gtreqless v^{BR}. \tag{9}$$

Proof of Lemma 4: We prove this by induction. The definition of v^{BR} is equivalent to

$$v^{BR}\left[\frac{1}{1-\delta^2} - 1\right] = \frac{c}{1-\delta^2} - g.$$

Hence we have that

$$[L(c_0, c_0) - L(c_1, c_0)]v^{BR} = L(c_0, c_0)v^I(c_0, c_0) - v^I(c_1, c_0)L(c_1, c_0).$$

It can be rewritten as

$$L(c_1, c_0)\{v^I(c_1, c_0) - v^{BR}\} = L(c_0, c_0)\{v^I(c_0, c_0) - v^{BR}\}.$$

Because $L(c_1, c_0) = 1 < L(c_0, c_0) = \frac{1}{1-\delta^2}$,

$$L(c_1, c_0)\{v^I(c_1, c_0) - v\} \gtreqless L(c_0, c_0)\{v^I(c_0, c_0) - v\} \iff v \gtreqless v^{BR},$$

and the assertion holds when $T = 0$.

Next suppose that the assertion holds for $T - 1$. We rewrite LHS inequalities for T as

$$\begin{aligned}
& L(c_{T+1}, c_T)\{v^I(c_{T+1}, c_T) - v\} \gtreqless L(c_T, c_T)\{v^I(c_T, c_T) - v\}, \\
\iff \quad & L(c_{T+1}, c_T)\{v^I(c_{T+1}, c_T) - v^{BR} - (v - v^{BR})\} \\
& \qquad \gtreqless L(c_T, c_T)\{v^I(c_T, c_T) - v^{BR} - (v - v^{BR})\}, \\
\iff \quad & \Gamma(c_{T+1}, c_T) - L(c_{T+1}, c_T)\{v - v^{BR}\} \gtreqless \Gamma(c_T, c_T) - L(c_T, c_T)\{v - v^{BR}\}.
\end{aligned}$$

By Lemma 3,

$$\begin{aligned}
\iff \quad & d - v^{BR} + \delta^2\Gamma(c_T, c_{T-1}) - \{1 + \delta^2 L(c_T, c_{T-1})\}\{v - v^{BR}\} \\
& \qquad \gtreqless d - v^{BR} + \delta^2\Gamma(c_{T-1}, c_{T-1}) - \{1 + \delta^2 L(c_{T-1}, c_{T-1})\}\{v - v^{BR}\} \\
\iff \quad & L(c_T, c_{T-1})\{v^I(c_T, c_{T-1}) - v\} \gtreqless L(c_{T-1}, c_{T-1})\{v^I(c_{T-1}, c_{T-1}) - v\},
\end{aligned}$$

and the last inequalities hold by the induction assumption. □

COROLLARY 1. *For any $T, T' \in \mathbb{N}$,*

$$\Gamma(c_T, c_T) = \Gamma(c_{T+1}, c_T).$$

COROLLARY 2. *$v^I(c_{T+1}, c_T) - v^I(c_T, c_T)$ is strictly decreasing in T.*

Proof of Corollary 2: In view of Corollary 1,

$$\begin{aligned}
v^I(c_{T+1}, c_T) - v^I(c_T, c_T) &= \frac{\Gamma(c_{T+1}, c_T)}{L(c_{T+1}, c_T)} - \frac{\Gamma(c_T, c_T)}{L(c_T, c_T)} \\
&= \frac{\Gamma(c_{T+1}, c_T)}{L(c_{T+1}, c_T)}\{1 - \frac{L(c_{T+1}, c_T)}{L(c_T, c_T)}\} \\
&= [v^I(c_{T+1}, c_T) - v^{BR}]\frac{L(c_T, c_T) - L(c_{T+1}, c_T)}{L(c_T, c_T)} \\
&= [v^I(c_{T+1}, c_T) - v^{BR}]\delta^{2(T+1)},
\end{aligned}$$

which is strictly decreasing in T. $\square$

Because of the concavity of $v(c_T; p_T^{T+1}(\alpha))$ and convexity of $v(c_{T+1}; p_T^{T+1}(\alpha))$, and continuity of average values with respect to T, the next lemma is immediate.

LEMMA 5. *For any $\delta > \delta^*$, there exists $0 \leq \tau^*(\delta) < \underline{\tau}(\delta)$ such that, if $\tau^*(\delta) < T < \underline{\tau}(\delta)$,*

(a) there exist $\underline{\alpha}_T^{T+1}(\delta) \in (0,1)$ and $\alpha_T^{T+1}(\delta) \in (0,1)$ with $\underline{\alpha}_T^{T+1}(\delta) < \alpha_T^{T+1}(\delta)$,

(b) $v(c_T; p_T^{T+1}(\alpha)) > v(c_{T+1}; p_T^{T+1}(\alpha)) \Leftrightarrow \alpha \in (\underline{\alpha}_T^{T+1}(\delta), \alpha_T^{T+1}(\delta))$.

Therefore, for sufficiently large T such that $\tau^*(\delta) < T < \underline{\tau}(\delta)$, there is a unique payoff-equalizing $\alpha_T^{T+1}(\delta)$. Let $\alpha_T^*(v^{BR})$ and $\alpha_{T+1}^*(v^{BR})$ be the fractions of c_T-strategy which solve $v(c_T; p_T^{T+1}(\alpha)) = v^{BR}$ and $v(c_{T+1}; p_T^{T+1}(\alpha)) = v^{BR}$ respectively. To show that the Best Reply Condition is satisfied at $\alpha_T^{T+1}(\delta)$, it suffices to prove

$$\alpha_{T+1}^*(v^{BR}) < \alpha_T^*(v^{BR}).$$

By computation, $v(c_T; p_T^{T+1}(\alpha)) = v^{BR}$ is equivalent to

$$v^I(c_T, c_T) - \frac{\{1-\alpha_T^*(v^{BR})\}L(c_T, c_{T+1})}{\alpha_T^*(v^{BR})L(c_T, c_T) + \{1-\alpha_T^*(v^{BR})\}L(c_T, c_{T+1})}\{v^I(c_T, c_T) - v^I(c_T, c_{T+1})\} = v^{BR}$$

$$\begin{aligned}
\iff \quad & \left[\alpha_T^*(v^{BR})L(c_T, c_T) + \{1-\alpha_T^*(v^{BR})\}L(c_T, c_{T+1})\right]\{v^I(c_T, c_T) - v^{BR}\} \\
& = \{1-\alpha_T^*(v^{BR})\}L(c_T, c_{T+1})\{v^I(c_T, c_T) - v^I(c_T, c_{T+1})\} \\
\iff \quad & \alpha_T^*(v^{BR}) = -\frac{\Gamma(c_T, c_{T+1})}{\Gamma(c_T, c_T) - \Gamma(c_T, c_{T+1})}.
\end{aligned}$$

Similarly, $v(c_{T+1}; p_T^{T+1}(\alpha)) = v^{BR}$ is equivalent to

$$\alpha_{T+1}^*(v^{BR}) = -\frac{\Gamma(c_{T+1}, c_{T+1})}{\Gamma(c_{T+1}, c_T) - \Gamma(c_{T+1}, c_{T+1})}.$$

Corollary 1 implies that

$$\begin{aligned}
& \{\alpha_T^*(v^{BR}) - \alpha_{T+1}^*(v^{BR})\}\{\Gamma(c_T, c_T) - \Gamma(c_T, c_{T+1})\}\{\Gamma(c_T, c_T) - \Gamma(c_{T+1}, c_{T+1})\} \\
= \; & \Gamma(c_T, c_T)\{\Gamma(c_{T+1}, c_{T+1}) - \Gamma(c_T, c_{T+1})\}.
\end{aligned}$$

Since $\Gamma(c_T, c_T) > 0$ for $T < \underline{T}(\delta)$, it suffices to prove that $\Gamma(c_{T+1}, c_{T+1}) > \Gamma(c_T, c_{T+1})$. In parameters,

$$\begin{aligned}\Gamma(c_{T+1}, c_{T+1}) &= \frac{1}{1-\delta^2}\left\{(1-\delta^{2(T+1)})d + \delta^{2(T+1)}c - v^{BR}\right\}\\ \Gamma(c_T, c_{T+1}) &= \frac{1-\delta^{2T}}{1-\delta^2}d + \delta^{2T}\ell - \frac{1-\delta^{2(T+1)}}{1-\delta^2}v^{BR}.\end{aligned}$$

Hence by computation,

$$\begin{aligned}&\left\{\Gamma(c_{T+1}, c_{T+1}) - \Gamma(c_T, c_{T+1})\right\}(1-\delta^2)\\ =\ & \delta^{2T}(1-\delta^2)(d-\ell) + \delta^{2(T+1)}(c - v^{BR}) > 0,\end{aligned}$$

because $(c - v^{BR})\delta^2 = (1-\delta^2)(g-c) > 0$. Therefore the Best Reply Condition is satisfied. Q.E.D.

REFERENCES

CARMICHAEL L., and B. MACLEOD (1997): "Gift Giving and the Evolution of Cooperation," *International Economic Review*, 38, 485-509.

DATTA, S. (1996): "Building Trust," Working Paper, London School of Economics.

ELLISON, G. (1994): "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching," *Review of Economic Studies*, 61, 567-588.

FUDENBERG, D., and E. MASKIN (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Review of Economics Studies*, 57, 533-556.

FUJIWARA-GREVE, T. (2002): "On Voluntary and Repeatable Partnerships under No Information Flow," *Proceedings of the 2002 North American Summer Meetings of the Econometric Society*. (http://www.dklevine.com/proceedings/game-theory.htm)

GHOSH, P., and D. RAY (1996): "Cooperation in Community Interaction without Information Flows," *Review of Economic Studies*, 63, 491-519.

KANDORI, M. (1992): "Social Norms and Community Enforcement," *Review of Economic Studies*, 59, 63-80.

KIM, Y-G., and J. SOBEL (1995): "An Evolutionary Approach to Pre-play Communication," *Econometrica*, 63, 1181-1193.

KRANTON, R. (1996a): "The Formation of Cooperative Relationships," *Journal of Law, Economics, & Organization* 12, 214-233.

——— (1996b): "Reciprocal Exchange: A Self-Sustaining System," *American Economic Review*, 86, 830-851.

MATSUI, A. (1991): "Cheap Talk and Cooperation in a Society," *Journal of Economic Theory*, 54, 245-258.

——— (1992): "Best Response Dynamics and Socially Stable Strategies," *Journal of Economic Theory*, 57(2), 343-362.

———, and M. OKUNO-FUJIWARA (2002): "Evolution and the Interaction of Conventions," *The Japanese Economic Review*, 53, 141-153.

MAYNARD SMITH, J. (1982) *Evolution and the Theory of Games*, Cambridge University Press.

OKUNO-FUJIWARA, M. (1987): "Monitoring Cost, Agency Relationship, and Equilibrium Modes of Labor Contract," *Journal of the Japanese and International Economies*, 1, 147-167.

———, and A. POSTLEWAITE (1995): "Social Norms and Random Matching Games," *Games and Econonomic Behavior*, 9, 79-109.

ROB, R., and H. YANG (2005): "Long-Term Relationships as Safeguards" mimeo., University of Pennsylvania.

SHAPIRO, C., and J. E. STIGLITZ (1984): "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*,74, 433-444.

WATSON, J. (2002): "Starting Small and Commitment," *Games and Economic Behavior*, 38, 176-199.