Voluntarily Separable Repeated Prisoner's Dilemma with Shared Belief*

Takako Fujiwara-Greve — Keio University, Dept. of Economics

Masahiro Okuno-Fujiwara — University of Tokyo, Graduate School of Economics

Nobue Suzuki — Komazawa University, Faculty of Economics

This version: March 31, 2009

Abstract: In Fujiwara-Greve and Okuno-Fujiwara (2009), an evolutionary stability concept was defined by allowing mutations of any strategy. However, in human societies, not all strategies are likely to be tried out when a player considers what happens in the future. In this paper we introduce the "shared belief" of potential continuation strategies, generated and passed on in a society, and mutations are restricted only among best responses against the shared belief. We show that a myopic strategy becomes a part of a bimorphic equilibrium under a shared belief and contributes to a higher payoff than ordinary neutrally stable distributions'.

Key words: shared belief, evolution, voluntary separation, Prisoner's Dilemma.
JLE classification: C 73.

*This work was supported by JSPS Grant-in-Aid for Scientific Research (C), No.19530158. The usual disclaimer applies.

1 Introduction

In Fujiwara-Greve and Okuno-Fujiwara (2009), henceforth abbreviated as Greve-Okuno, a model of voluntarily separable repeated Prisoner's Dilemma (VSRPD) is formulated and analyzed with neutrally stable distribution (NSD) as an associated solution concept. In this paper, we argue that we need to introduce a modified solution concept to appropriately analyze VSRPD in a human society and provide further analyses of VSRPD with this modified concept.

A VSRPD consists of a matching process, in which pairs of players are ramdomly formed, and a component game, which is a repeatable Prisoner's Dilemma on a voluntary basis[1]. In each period, players without a partner are randomly matched and play the Prisoner's Dilemma (see Table 1). The matched players observe the action profile of only themselves (i.e., not of players of other matches). At the end of a period, partners simultaneously choose whether to keep the partnership or to put an end to it based on the observations. If at least one partner chooses to end the partnership, both playeres become unmatched and return to the random matching process in the following period. There is also a possibility of death of a partner which forces the surviving player (and a new born player who replaces the deceased) to go back to the matching pool. If both partners choose to keep the partnership and survive the period, the same partners play the Prisoner's Dilemma in the next period, skipping the matching process. The game coninues this way.

The VSRPD framework is developed to illustrate the contemporary society where matching is rather easy, through improved communication and matching channels via Internet and other globalized networks, but it is also easy to run away without carrying the reputation. As compared to ordinary repeated games where the fixed set of players play the game throughout and usual random matching games where the players must be re-matched every period, the VSRPD covers an intermediate case and can encompass both classes of games by strategic choice of

P1 \ P2	C	D
C	c, c	ℓ, g
D	g, ℓ	d, d

Table 1: Prisoner's Dilemma ($g > c > d > \ell$)

[1] Thus, the component game is potentially an infinite period extensive-form game.

players. Both repeated games and random matching games must rely on personal histories to enforce cooperation in the Prisoner's Dilemma.[2] In VSRPD, however, even if no personal history of a newly matched player is known, gradual build-up of cooperation may emerge as an equilibrium behavior and, if so, a threat of breaking up the partnership and subsequent reduction of continuation payoff provides an incentive to sustained cooperation.

Greve-Okuno (2009) employed *neutrally stable distribution* as the solution concept. A mutant strategy is said to *invade* the current strategy distribution if, when some positive fraction of current distribution is replaced by this mutant strategy, the mutant's lifetime payoff is strictly higher than that of some incumbent strategy. A strategy distribution is neutrally stable if no new mutant strategy can invade the current strategy distribution.

To illustrate the neutral stability concept, consider the *hit-and-run* strategy, written as the d_0-strategy, which plays D in the first period of any match and ends the partnership regardless of the first period outcome of the Prisoner's Dilemma. Clearly, the strategy distribution consisting only of d_0-strategy (the d_0-monomorphic distribution) is a Nash equilibrium, since any strategy must play a one-shot Prisoner's Dilemma against d_0-strategy and hence no strategy can earn a higher payoff than d_0-strategy does.

However, the d_0-monomorphic distribution is not neutrally stable. As Robson (1990) and Matsui (1991) showed in a pre-play cheap talk model, a strategy that imitates the d_0-strategy but cooperates within themselves can invade the population. For example, consider the (one-period) *trust-building* strategy, written as the c_1-strategy, which plays D in the first period and keeps the partnership if and only if the first period outcome is (D, D). If the partnership continues to period $t \geq 2$, the strategy plays C in the Prisoner's Dilemma and keeps the partnership if and only if the current outcome is (C, C). When matched with d_0, the c_1-strategy obtains per period payoff of d, the same payoff level as when the d_0-strategy is matched with another d_0. However, when matched with c_1, the c_1-strategy earns more than d on average, while d_0-strategy always earns d against any strategy. Thus, the c_1-strategy invades the d_0-monomorphic distribution.

Although the above story fits genetically programmed organisms, which mechanically plays a mutant strategy without paying attention to the partner's possible reactions, it does not seem to fit humans well, who are forward-looking and conscious about social environment. Applying

[2]See for example Fudenberg and Maskin (1986) for repeated games, and Kandori (1992) for the random matching game with Prisoner's Dilemma as the stage game.

evolutionary game theory to human society, we usually think (e.g., Kandori, Mailath and Rob (1993)) introduction of a mutation strategy as an intentional trial by *experiment*. Note that our component game is an extensive form game with a strategy as a plan of contingent actions. In such a case, even if a human player starts to play a new (mutant) strategy as an experiment, she is likely to evaluate its consequences (by forecasting the partner's future play) before continuing its use. Let us consider the above story in such contexts.

Upon infiltrating into the d_0-monomorphic distribution, what mutant c_1 is supposed to do is to play k (keep) after observing (D, D) and to play C if the partner also plays k. Notice, however, that the history of $((D, D), (k, k))$ should not have occurred under the incumbent distribution. Upon observing this novel history, a forward-looking player may not have confidence that the partner will necessarily play C. Alternatively, she may expect that the partner may play D in the second period if the partnership continues. If a human player is to ponder which possibility is more likely, then we must consider the belief over the off-path plays. At the history $((D, D), (k, k))$, both partners are taking mutant strategies and thus there is nothing reliable that helps one predict the partner's next action choice. We argue that in such a case one is likely to resort to the common sense such as teachings from one's parents or social norm. We call these common senses as "shared belief" within the society, which are reproduced over generations and maintained within ths society over time. Thus, even when one is trying out a new strategy, one must have the shared belief in mind.

For example, given $((D, D), (k, k))$, suppose it is the shared belief that "any player is going to play C from the second period on". In this case, it is rational to play C and enjoy the payoff of c for $t \geq 2$. That is, under this belief, you expect that your partner will use the c_1-strategy, to which playing c_1 is a best response. It follows that, the c_1-strategy invades the d_0-monomorphic distribution under this shared belief. Suppose, instead, that the belief that "any player is going to play D after $((D, D), (k, k))$" is shared in the society. Then it is to your interest to play D, rather than to play C. It follows that c_1 cannot invade the d_0-monomorphic distribution and it is stable under this second belief. In sum, the d_0-monomorphic distribution can be both stable and unstable, depending on the belief shared within the society.

By incorporating the shared belief in the stability, we can extend the analysis of Greve-Okuno (2009) in two directions: a new *bimorphic* equilibrium consisting of two strategies exists, and it is not only more efficient than the one found in Greve-Okuno (2009) but also more realistic. In Greve-Okuno (2009), it was shown that for some parameter values, there exists a bimorphic neutrally stable distribution consisting of the c_1-strategy and the c_0-strategy, which starts any partnership with C and keep the partnership if and only if the current outcome is (C, C). Notice that any bimorphic distribution of the c_0-strategy and the d_0-strategy is not neutrally stable since the c_1-strategy can invade it. However, the introduction of the concept of shared belief changes the picture entirely. With the shared belief that "D will be played after $((D, D), (k, k))$", the c_1-strategy cannot invade the bimorphic Nash equilibrium of c_0 and d_0. Moreover, the interpretation of the bimorphic equilibrium of c_0 and d_0 is easier than that of c_0 and c_1. The former can be interpreted as co-existence of inherently cooperative players and inherently defective players, which is often assumed in incomplete information models, while the latter does not have inherently defective players.

This paper is organized as follows. In Section 2 we describe the model and basic stability concepts such as Nash equilibrium and neutral stability. In Section 3 we define the stability under the shared belief and do a preliminary analysis of monomorphic distributions. In Section 4 we show that a bimorphic distribution of c_0 and d_0 is stable under the shared belief and has various good properties such as existence and efficiency. Section 5 gives concluding remarks.

2 Model and Basic Stability Concepts

2.1 VSRPD Model

Consider a large society of a continuum of players of measure 1. The time is discrete. In each period, players without a partner enter the random matching pool and form pairs to play the *Voluntarily Separable Repeated Prisoner's Dilemma* (VSRPD) as follows.

Newly matched players have no knowledge of the past action history of each other, and they play the Prisoner's Dilemma (see Table 1) by choosing action C or D simultaneously. The actions in the Prisoner's Dilemma are observable only by the partners. At the end of a period, the partners simultaneously choose whether to keep the partnership (action k) or end it (action e), based only on the observations. After the continuation decisions, at the end of a period, each player may exit from the society for some exogenous reason (which we call a "death") with probability $1 - \delta$ where $0 < \delta < 1$. If a player dies, a new player enters into the society, keeping

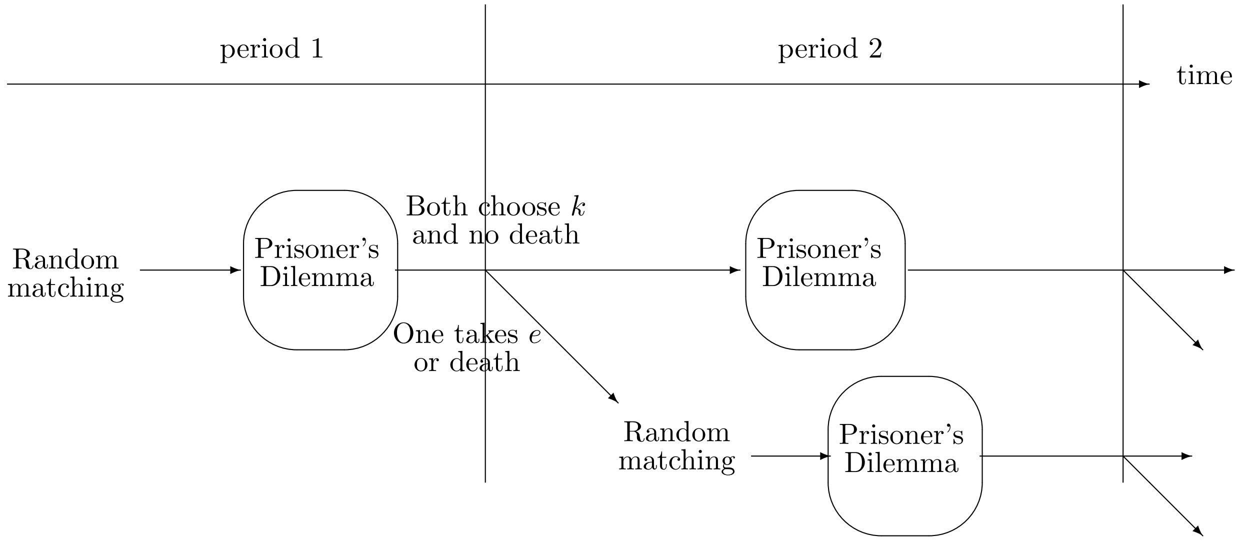

Figure 1: Outline of the VSRPD

the population size constant.

The partnerships continue/dissolve across periods as follows. If at least one player chooses e, the partnership ends and both players (if they survive) go to the random matching pool in the next period. If both chose action k, unless one of them dies, they stay as partners and play the Prisoner's Dilemma in the next period, skipping the matching process. If the partner dies, the surviving player and a newborn player who replaced the deceased go to the random matching pool in the next period. The outline of VSRPD is depicted in Figure 1.

The one-shot payoffs in the Prisoner's Dilemma are shown in the Table 1. Assume that $g > c > d > \ell$ and $2c \geqq g + \ell$. The latter is for simplicity and to make the symmetric action profile (C, C) efficient.

The game continues with probability δ from an individual player's point of view. Thus we focus on the expected total payoff, with δ being the effective discount factor of a player.

Let $t = 1, 2, \ldots$ indicate the periods in a match, not the calendar time in the game. Under the no-information-flow assumption, we focus on match-independent strategies that only depend on t and the private history of actions within a match. Let $H_t := [\{C, D\}^2 \times \{k, e\}^2]^{t-1}$ be *the set of partnership histories at the beginning of* $t \geqq 2$ and let $H_1 := \{\emptyset\}$.

Definition: A *pure strategy* s of VSRPD consists of $(x_t, y_t)_{t=1}^{\infty}$ where:

- $x_t : H_t \to \{C, D\}$ specifies an action choice $x_t(h_t) \in \{C, D\}$ given the partnership history[3]

[3]Note that only (k, k) throughout the past would allow players to choose actions.

$h_t \in H_t$, and

- $y_t : H_t \times \{C, D\}^2 \to \{k, e\}$ specifies whether to keep or end the partnership, depending on the partnership history $h_t \in H_t$ and the current period action profile.

The set of pure strategies of VSRPD is denoted as $\boldsymbol{S}$ and the set of all strategy distributions in the population is denoted as $\mathcal{P}(\boldsymbol{S})$. We assume that each player uses a pure strategy, which is natural in an evolutionary game and simplifies the analysis.

We investigate the evolutionary 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 is also stationary, thanks to the stationary death process.[4]

2.2 Average and Lifetime Payoffs

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 s' intend to maintain the match, it will only continue with probability δ^2. Suppose that the planned length of the partnership of s and s' is $T(s, s')$ periods, if no death occurs. 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 $V(s, s')$. The*average per period payoff* that s expects to receive within the match with s' is denoted as $v(s, s')$. Clearly,

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

Next we show the structure of the lifetime and average payoff of a player endowed with strategy $s \in \boldsymbol{S}$ in the matching pool, waiting to be matched randomly with a partner. When a strategy distribution in the matching pool is $p \in \mathcal{P}(\boldsymbol{S})$ and is stationary, we write the *expected total discounted value of payoff streams* s expects to receive during his lifetime as $V(s; p)$ and the average per period payoff s expects to receive during his lifetime as

$$v(s; p) := \frac{V(s; p)}{L} = (1 - \delta)V(s; p),$$

[4]See Greve-Okuno (2009) footnote 7 for details.

where $L = 1 + \delta + \delta^2 + \cdots = \frac{1}{1-\delta}$ is the expected lifetime of s.

Thanks to the stationary distribution in the matching pool, we can write $V(s;p)$ as a recursive equation. If p has a finite/countable support, then we can write

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

where $supp(p)$ is the support of the distribution p, 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 $T(s,s')$ to go back to the matching pool. Thanks to the stationarity of p, the continuation payoff after a match ends for any reason is always $V(s;p)$.

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

$$\begin{aligned} V(s;p) &= \sum_{s' \in supp(p)} p(s') \Big[V(s,s') + \{1 - (1-\delta)L(s,s')\} V(s;p) \Big] \\ &= \sum_{s' \in supp(p)} p(s') V(s,s') + \{1 - \frac{L(s;p)}{L}\} V(s;p). \end{aligned} \tag{2}$$

Hence the average payoff is a nonlinear function of the strategy distribution p:

$$v(s;p) = \frac{V(s;p)}{L} = \sum_{s' \in supp(p)} p(s') \frac{L(s,s')}{L(s;p)} v(s,s'), \tag{3}$$

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'. This nonlinearity is the important characteristics of the voluntarily separable game and is due to the endogenous duration of partnerships. Note also that, if p is a strategy distribution consisting of a single strategy s', then $v(s;p) = v(s,s')$.

2.3 Nash Equilibrium and Neutrally Stable Distribution

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 $s' \in \boldsymbol{S}$,

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

and is deoted 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 $s \in \text{supp}(p)$, $s \in BR(p)$.

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 $\epsilon \in (0, \bar{\epsilon}_q)$, $Eu(p, (1-\epsilon)p + \epsilon q) \geqq Eu(q, (1-\epsilon)p + \epsilon q)$. (Maynard Smith, 1982.)

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

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 for any $s \in supp(p)$ and any $\epsilon \in (0, \bar{\epsilon})$,

$$v(s; (1-\epsilon)p + \epsilon p_{s'}) \geqq v(s'; (1-\epsilon)p + \epsilon p_{s'}), \tag{4}$$

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

If a *monomorphic* distribution consisting of a single strategy constitutes a NSD, the strategy is called a Neutrally Stable Strategy (NSS). It can be easily seen that any NSD is a Nash equilibrium.

We note that in VSRPD, there is no Nash equilibrium in which all players play C in the first period of a match. This was shown in Greve-Okuno (2009).[5]

Lemma 1. *(Greve-Okuno, 2009, Lemma 1) Any strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$ such that all strategies in the support start with C in $t = 1$ is not a Nash equilibrium.*

The intuition is that if all players play C in the first period of a partnership, then a strategy that chooses D in the first period and then runs away can earn a higher payoff.

3 Monomorphic Strategy Distribution and Stability under the Shared Belief

The Nash equilibrium and neutral stability are based only on the payoff comparison at the beginning of the game. If we consider mutations as trial adoptions of new strategies by human

[5]From now on, if a result is due to Greve-Okuno (2009), we make that clear and refer the readers to that paper for proofs.

players, however, it is also natural to compare continuation payoffs at all histories during the game before adopting a new strategy. We now describe how stability of a strategy (or a strategy distribution) may change with or without considerations of the continuation payoffs, based on beliefs over the continuation strategy of the partner.

Definition: Let the *strict myopic strategy* (written as d_0-strategy) be a strategy as follows: Play D and end the partnership after any partnership history, including the first period.

It is easy to see that for any parameter combination (δ, g, c, d, ℓ), the strategy distribution consisting only of the strict myopic strategy d_0 is a Nash equilibrium. However, it is not neutrally stable.

Lemma 2. *(Greve-Okuno, 2009, Lemma 2) For any parameter combination* (δ, g, c, d, ℓ)*, the strict myopic strategy* d_0 *is not a NSS.*

This is because the monomorphic distribution of the d_0-strategy can be invaded by the following "one-period trust-building" strategy (written as c_1-strategy):

- $t = 1$: Play D and keep the partnership if and only if (D, D) is observed in this period.
- $t \geqq 2$: Play C and keep the partnership if and only if (C, C) is observed in the current period.

The c_1-strategy earns the same average payoff d in the population of the d_0-strategy as the d_0-strategy does and earns more than d if it meets the same c_1-strategy, which occurs with a small but positive probability. Hence, its average payoff strictly exceeds that of the incumbent d_0-strategy.

However, one can question the mutation to C after observing $((D, D), (k, k))$ in the first period, when the incumbent population consists only of the strict myopic strategy d_0. In this population, it is natural to suppose that the shared belief of the continuation strategy after any history is again the d_0-strategy[6]. Then playing C is never a best response to the shared belief, and it is unlikely that such mutation occurs. Thus we can say that the distribution of the d_0-strategy with the shared belief being d_0 after any history cannot be invaded by the c_1-strategy.

[6]Note that in VSRPD, at the beginning of any period, the set of continuation strategies coincides with the set of strategies, and thus we can call a continuation strategy by the same name as a strategy.

Consider a slightly different myopic strategy.

Definition: Let the *reconcilable myopic strategy* (written as $\tilde{d}_0$-strategy) be a strategy as follows: Play D and end the partnership in the first period of a match. If the partnership continued to $t \geqq 2$, then play C and keep if and only if the current period action profile is (C, C).

Clearly, the play path of the monomorphic distribution of the $\tilde{d}_0$-strategy is the same as that of the monomorphic distribution of the d_0-strategy. However, unlike the case of the d_0-strategy, the natural shared belief under the $\tilde{d}_0$-monomorphic distribution is that the partner plays C after $((D, D), (k, k))$. Then the mutation to the c_1-strategy is a best response to the shared belief after $((D, D), (k, k))$, and thus the c_1-strategy can invade this population.

In sum, depending on the shared belief over the off-path plays, the scope of mutations can be restricted. This idea is related to properness such that certain mutations are more likely than others (Myerson, 1978) and the likelihood is determined by sequential rationality considerations (Kreps and Wilson, 1982). Note also that this distinction of mutations is relevant since our component game is an extensive form game.

We now formalize the above argument.

Definition: For any $t = 2, 3, \ldots$, any partnership history h_t at the beginning of t and a strategy $s = (x_t, y_t)_{t=1}^{\infty}$, the *continuation strategy* (within one partnership) of s given h_t, written as $s|_{h_t}$, consists of $(x_n, y_n)_{n=1}^{\infty}$ such that

- for any history $h_n \in H_n$ with the same partner, $x_n(h_n) = x_{t+n}(h\, h_n)$, where $(h\, h_n) \in H_{t+n}$ is the concatination of the two histories;
- for any $(h_n, a(t+n+1)) \in H_n \times \{C, D\}^2$ with the same partner, $y_n(h_n, a(t+n+1)) = y_{t+n}(h\, h_n, a(t+n+1))$.

The continuation payoff of a strategy s after a partnership history h when the current partner is expected to play a continuation strategy $\hat{s}|_h$ and the stationary strategy distribution in the matching pool is p is defined as follows.

$$V(s|_h; \hat{s}|_h, p) = V(s|_h, \hat{s}|_h) + \{1 - (1-\delta)L(s|_h, \hat{s}|_h)\}V(s; p).$$

The average continuation payoff is defined as $v(s|_h; \hat{s}|_h, p) = (1-\delta)V(s|_h; \hat{s}|_h, p)$. Using this, given a stationary strategy distribution in the matching pool, if the belief of the partner's continuation strategy is given, the best response to it is well-defined.

In this paper we focus on strategy distributions such that there naturally exists a unique shared belief regarding one's partner's continuation strategy at any off-path information set, as d_0-strategy and $\tilde{d}_0$-strategy. Among those, we select stable distributions which cannot be invaded by mutations that are best responses to the shared belief. A more general formulation of the shared belief and stability under the shared belief is left for a near future research.

Definition: Given a stationary strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$, the set of *on-path histories* $H(p)$ and the set of *off-path histories* $H^c(p)$ are defined as follows:

$$\begin{aligned} H(p) &:= \{h \in H \mid \exists s, s' \in supp(p) \text{ such that } h \text{ occurs with a positive probability} \\ &\qquad \text{when } s \text{ and } s' \text{ are matched}\}, \\ H^c(p) &:= H \setminus H(p). \end{aligned}$$

Definition: A stationary strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$ *has a shared belief* $\mu(p) \in \boldsymbol{S}$ if, for any off-path history $h \in H^c(p)$ and any $s, s' \in supp(p)$,

(i) $s\mid_h = s'\mid_h = \mu(p)$; and

(ii) $s\mid_h$ is a best response to $\mu(p)$.

We thus focus on strategy distributions with a unique continuation strategy by the partner at any off-path history, which becomes the shared belief, and the shared belief is passed on because it is a best response to itself. The above definition can be weakened to allow off-path history dependent shared belief, i.e., $\mu(p)$ can be dependent on $h \in H^c(p)$, but in this paper we focus on monomorphic and bimorphic distributions with a unique off-path continuation strategy.

Definition: A stationary strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$ is *stable under shared belief* (SSB) if

(i) p has a shared belief $\mu(p) \in \boldsymbol{S}$; and

(ii) and no strategy that is a best response to the shared belief after some off-path history $h \in H^c(p)$ can invade the population, i.e., for any $z \in \boldsymbol{S}$,

(ii-a) either $z\mid_h$ is not a best response to $\mu(p)$ for some $h \in H^c(p)$, or (ii-b) there exists $\bar{\epsilon} > 0$ such that for any $\epsilon \in (0, \bar{\epsilon})$, $v(s; (1-\epsilon)p + \epsilon p_z) \geqq v(z; (1-\epsilon)p + \epsilon p_z)$ for all $s \in supp(p)$.

From this definition it is clear that if a strategy distribution is a NSD then it is stable under the shared belief. To formalize the stability of the monomorphic distributions of myopic strategies, we give an additional definition.

Definition: Let $\tilde{c}_0$-strategy be as follows:
In any $t = 1, 2, \ldots$, play C and k if and only if the current action profile is (C, C).

Then the shared belief under the reconcilable myopic strategy is $\tilde{c}_0$, while the shared belief under the strict myopic strategy is d_0.

Proposition 1. *For any parameter combination (δ, g, c, d, ℓ), the monomorphic strategy distribution consisting only of the strict myopic strategy d_0 is SSB. For sufficiently high δ, the monomorphic distribution consisting only of the reconcilable myopic strategy $\tilde{d}_0$ is not SSB.*

Proof: Take the monomorphic distribution of the d_0-strategy. For any $z \in \boldsymbol{S}$, if z plays differently on the play path against d_0, then it earns less than the d_0-strategy does. Hence it suffices to consider z that plays different off the play path against d_0.

For any history $h \in H^c(d_0)$, the shared belief is again d_0. The relevant mutant strategies that are best responses to d_0 are those that play D after h, and therefore no such strategy can earn higher average lifetime payoff than the d_0-strategy does, even if they meet with each other. Therefore the monomorphic distribution of the d_0-strategy is stable under the shared belief.

Next, take the monomorphic distribution of the $\tilde{d}_0$-strategy. Again it suffices to consider $z \in \boldsymbol{S}$ that plays differently off the play path against $\tilde{d}_0$. After any off-path history $h \in H^c(\tilde{d}_0)$, either c_1-strategy or a strategy that plays D is a best response to the continuation strategy of $\tilde{d}_0$. Since both $\tilde{d}_0 \mid_h$ and $c_1 \mid_h$ are the same as the $\tilde{c}_0$-strategy,

$$V(c_1|_h; \tilde{d}_0|_h, p_{\tilde{d}_0}) = \frac{c}{1-\delta^2} + \frac{\delta d}{1-\delta^2}.$$

Consider a one-step deviation to play D after an off-path history h. This gives the continuation payoff of

$$g + \delta \frac{d}{1-\delta}.$$

By computation,

$$\frac{c}{1-\delta^2} + \frac{\delta d}{1-\delta^2} \geqq g + \delta \frac{d}{1-\delta} \iff \delta^2 \geqq \frac{g-c}{g-d}. \tag{5}$$

Therefore for sufficiently large δ, c_1-strategy is a best response to the shared belief. Moreover c_1-strategy earns on average higher than d when meeting the same strategy, thus for any $\epsilon > 0$, $v(c_1; (1-\epsilon)p_{\tilde{d}_0} + \epsilon p_{c_1}) > v(\tilde{d}_0; (1-\epsilon)p_{\tilde{d}_0} + \epsilon p_{c_1})$. □

Proposition 1 shows that the stability under the shared belief selects among Nash equilibria. However, the stable monomorphic distribution of the d_0-strategy earns the average payoff of only d. If there is a monomorphic distribution that plays (C, C) for some periods, that is more efficient. By Lemma 1, among monomorphic strategy distributions, the c_1-strategy is the most efficient one since it plays (C, C) forever after the first period of a match. The following condition is sufficient for the c_1-strategy to be a NSS and hence stable under the shared belief.

Proposition 2. *For any* (δ, g, c, d, ℓ) *such that* $g + d < 2c$ *and* $\delta^4 > \max\left\{\frac{(g-c)^2}{(c-d)^2}, \frac{(g-\ell)-(c-d)}{g-\ell}\right\}$, *the* c_1*-strategy is a NSS.*

Proof: See Appendix.

Hence if the above condition is satisfied, then the monomorphic SSB distribution of the d_0-strategy is an inefficient equilibrium.

4 Bimorphic Strategy Distributions

One of the contributions of Greve-Okuno (2009) is that for sufficiently large δ and for some payoff combinations (g, c, d, ℓ), there exists a *bimorphic* neutrally stable distribution, consisting of two strategies, which is more efficient than any monomorphic NSD. Specifically, in some bimorphic NSD, some players can play C in the first period of a match. Then the average payoff is greater than any monomorphic NSD, since no monomorphic strategy distribution can start with C by Lemma 1.

We first give a sufficient condition for the existence of a bimorphic NSD consisting of the c_1-strategy and the following c_0-strategy (which is slightly different from the definition in Greve-Okuno, 2009, to induce the unique shared belief).

Definition: Let the c_0-strategy be as follows:

In any $t = 1, 2, \ldots$, play C and k if the partnership history is either empty or consists only of $((C, C), (k, k))$. Play D and e otherwise.

Denote the two-strategy distribution of c_0 and c_1 such that $\alpha \in (0, 1)$ is the fraction of c_0 by

$$p_0^1(\alpha) = \alpha p_{c_0} + (1 - \alpha) p_{c_1}.$$

If $p_0^1(\alpha^N)$ for some $\alpha^N \in (0, 1)$ is neutrally stable, then it is more efficient than any monomorphic NSD.

Proposition 3. *Let*

$$\begin{aligned} D_{01} &= \Big\{(1 + 2\delta^2)c + (1 - 2\delta^2)d - (1 + \delta^2)\ell - (1 - \delta^2)g\Big\}^2 \\ &\quad -4\delta^2(c - \ell + g - d)\{\delta^2(c - d) + (d - \ell)\}. \end{aligned}$$

There exists $\alpha^N \in (0, 1)$ such that the bimorphic distirbution $p_0^1(\alpha^N)$ is a NSD, if $D_{01} > 0$ and one of the following conditions hold:

(a) $c + d \geqq g + \ell$ and $\delta^2 \geqq \frac{c+d-g-\ell}{g-\ell}$;

(b) $c + d < g + \ell$ and $\delta^2 \geqq \frac{-(c+d-g-\ell)}{2(c-d)+g-\ell}$.

Proof: See Appendix.

Note that the above condition does not converge to the set of all parameter combinations (g, c, d, ℓ) even if δ converges to 1, since the condition $D_{01} > 0$ is still binding.

By contrast, the bimorphic distribution with the support $\{c_0, d_0\}$ is not a NSD. This is because c_1-strategy can invade the distribution, by the same logic as Lemma 2. However, as the d_0-strategy became stable under the shared belief, we can show that the bimorphic Nash equilibrium of c_0 and d_0-strategy is stable under the shared belief.

Proposition 4. *For suffciently large δ, there exists $\alpha^S \in (0, 1)$ such that the bimorphic distribution $\alpha^S p_{c0} + (1 - \alpha^S) p_{d0}$ is SSB.*

Proof: See Appendix.

Note that the above proposition includes that as δ converges to 1, any payoff parameter combination (g, c, d, ℓ) admits the bimorphic SSB distribution.

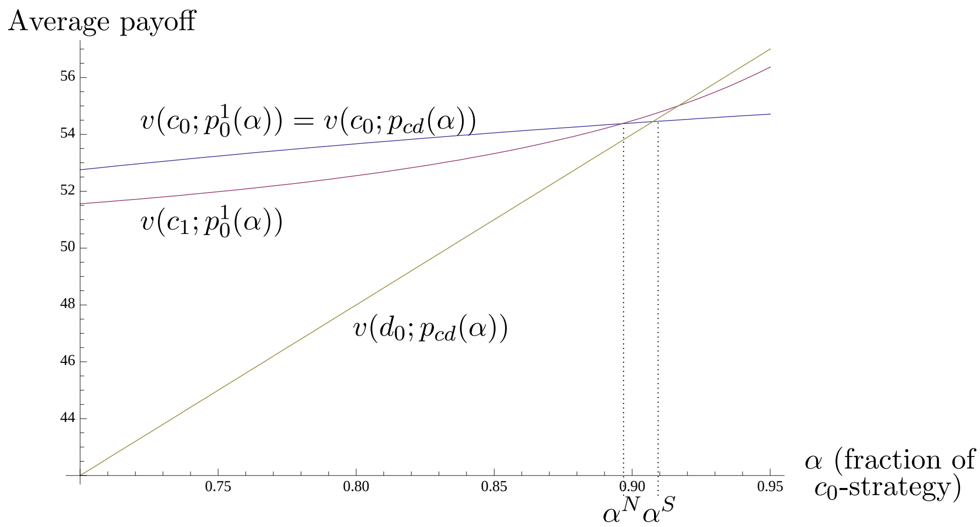

Figure 2: Bimorphic Equilibria

Under the same payoff parameters, the bimorphic NSD consisting of the c_0-strategy and the c_1-strategy gives strictly smaller average payoff than the bimorphic SSB distribution consisting of c_0-strategy and d_0-strategy. This is because the latter has more share of the "good" type c_0-strategy. See Figure 2. Therefore, the SSB distribution $\alpha^S p_{c0} + (1 - \alpha^S) p_{d0}$ is better than the NSD $\alpha^N p_{c0} + (1 - \alpha^N) p_{c1}$ in two aspects: it exists for general payoff parameter combinations and it is more efficient.

Proposition 5. *For any parameter combination (δ, g, c, d, ℓ) such that α^N and α^S exist simultaneously, $v(c_0; p_0^1(\alpha^N)) < v(c_0; p_{cd}(\alpha^S))$.*

Proof: See Appendix.

Using the general concept of *trust-building strategy* in Greve-Okuno (2009), which is a generalization of the c_1-strategy, we can strengthen the efficiency result.

Definition: For any $T = 0, 1, 2, \ldots$, let a *trust-building strategy* with T periods of trust-building (written as c_T-strategy) be a strategy as follows:

- $t \leqq T$: Play D and keep the partnership if and only if (D, D) is observed in the current period.
- $t \geqq T + 1$: 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 the *trust-building phase* and the periods afterwards are called the *cooperation phase*. A trust-building strategy continues the partnership if and only if an "acceptable" action profile is played. The acceptable action profile during the trust-building phase is (D, D) only, and during the cooperation phase, it is (C, C) only.

Corollary 1. *The SSB bimorphic distribution* $\alpha^S p_{c0} + (1 - \alpha^S) p_{d0}$ *is more efficient than any bimorphic distribution consisting of two trust-building strategies of the form* $\alpha p_{cT} + (1-\alpha) p_{cT+1}$.

Therefore, by introducing the shared belief, we could improve the efficiency under bimorphic distributions. This is a contrast to the efficiency comparison among monomorphic distributions.

5 Concluding Remarks

The co-existence of two trust-building strategies in Greve-Okuno (2009) can be interpreted as a "mis-match" of their trust-building phase length. In this paper we showed the co-existence of an out-for-tat strategy c_0 and the strict myopic strategy d_0, which can be interpreted as inherently cooperative players and inherently defective players. The latter is often assumed in the incomplete information games. Thus we provided an evolutionary foundation to such models (e.g., Ghosh and Ray, 1996, and Kranton, 1996).

Moreover, the bimorphic distribution of an out-for-tat strategy c_0 and the strict myopic strategy d_0 is more efficient than any bimorphic distribution of trust-building strategies. Hence if we introduce group selection, it is likely that the former will dominate the latter in the long run.

APPENDIX: Proofs

Proof of Proposition 2: Following Greve-Okuno (2009), consider one-step deviation in $t \geqq 2$ to play D. (Deviations in $t = 1$ or in the continuation decision nodes are not relevant.) Then the partner playing c_1-strategy will end the partnership so that the deviant strategy receives

$$g + \delta V(c_1; p_1),$$

where p_1 is the monomorphic distribution of the c_1-strategy. From (1), the payoff of the c_1-strategy from the point of Prisoner's Dilemma action choice in $t \geqq 2$ is

$$V(c_1, c_1) + \delta(1-\delta)\{1 + \delta^2 + \delta^4 + \cdots\}V(c_1; p_1) = \frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2}V(c_1; p_1).$$

Thus the monomorphic distribution of the c_1-strategy is a Nash equilibrium if and only if

$$\begin{aligned} & g + \delta V(c_1; p_1) \leqq \frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2}V(c_1; p_1) \\ \iff \quad & \delta^2(1-\delta)V(c_1; p_1) \leqq c - (1-\delta^2)g \\ \iff \quad & v(c_1; p_1) \leqq \frac{1}{\delta^2}\{c - (1-\delta^2)g\} =: v^{BR}. \end{aligned} \tag{6}$$

Since

$$v(c_1; p_1) = \frac{d + \frac{\delta^2 c}{1-\delta^2}}{1/(1-\delta^2)} = (1-\delta^2)d + \delta^2 c,$$

(6) is satisfied if and only if

$$\begin{aligned} & \frac{1}{\delta^2}\{c - (1-\delta^2)g\} \geqq (1-\delta^2)d + \delta^2 c \\ \iff \quad & (1-\delta^4)c - (1-\delta^2)g \geqq \delta^2(1-\delta^2)d \\ \iff \quad & (1+\delta^2)c - g \geqq \delta^2 d \\ \iff \quad & \delta^2 \geqq \frac{g-c}{c-d}. \end{aligned}$$

To warrant that such δ exists within $(0, 1)$, we need $c - d > g - c$, that is $2c > g + d$.

In addition, the condition (6) of Greve-Okuno (2009) must be satisfied for the neutral stability:

$$(1-\delta^4)(g-\ell) < c - d.$$

Therefore, the overall sufficient condition is

$$\delta^4 > \max\left\{\frac{(g-c)^2}{(c-d)^2}, \frac{(g-\ell)-(c-d)}{g-\ell}\right\}.$$

□

Proof of Proposition 3: The average payoff of the two strategies are as follows.

$$\begin{aligned} v(c_0; p_0^1(\alpha)) &= \frac{\alpha\frac{c}{1-\delta^2} + (1-\alpha)\ell}{\alpha\frac{1}{1-\delta^2} + (1-\alpha)}, \\ v(c_1; p_0^1(\alpha)) &= \frac{\alpha g + (1-\alpha)(d + \delta^2\frac{c}{1-\delta^2})}{\alpha + (1-\alpha)\frac{1}{1-\delta^2}}. \end{aligned}$$

Therefore the two strategies have the same average payoff if and only if

$$\begin{aligned} & \{v(c_0; p_0^1(\alpha)) - v(c_1; p_0^1(\alpha))\}\{\alpha \frac{1}{1-\delta^2} + (1-\alpha)\}\{\alpha + (1-\alpha)\frac{1}{1-\delta^2}\}(1-\delta^2) = 0; \\ \iff \quad & f(\alpha) := A\alpha^2 + B\alpha + C = 0, \end{aligned}$$

where

$$\begin{aligned} A &= -\delta^2(c - \ell + g - d) < 0; \\ B &= (1+2\delta^2)c + (1-2\delta^2)d - (1+\delta^2)\ell - (1-\delta^2)g \\ C &= -\{\delta^2(c-d) + (d-\ell)\} < 0. \end{aligned}$$

Thus the quadratic function $f(\alpha)$ should have a real solution α^N within $(0,1)$ such that for α in a neighborhood of α^N,

$$\alpha \gtreqless \alpha^N \iff 0 \gtreqless f(\alpha) = A\alpha^2 + B\alpha + C.$$

Since $A < 0$, f is a concave function of α. By computation $f(1) = A + B + C = c - g < 0$ and $f(0) = C < 0$, and thus we need the discriminant $D_{01} := B^2 - 4AC$ to be positive as well as $f'(0) > 0$ and $f'(1) < 0$.

By computation,

$$\begin{aligned} f'(0) > 0 \quad &\iff \quad B = c + d - \ell - g + \delta^2(2c - 2d - \ell + g) > 0; \\ f'(1) < 0 \quad &\iff \quad 2A + B = c + d - \ell - g - \delta^2(g - \ell) < 0. \end{aligned}$$

Thus we can divide into two cases. If $c + d - \ell - g \geqq 0$, then $f'(0) > 0$ for any $\delta \in (0,1)$ so that $f'(1) < 0$ is binding. In this case we need $\delta^2 \geqq (c + d - \ell - g)/(g - \ell)$.

If $c + d - \ell - g < 0$, then $f'(1) < 0$ for any δ so that we need $\delta^2 \geqq -(c + d - \ell - g)/(2c - 2d - \ell + g)$. □

Proof of Proposition 4: It is straightforward by following the argument in Proposition 1 to show that a Nash equilibrium $ap_{c_0} + (1-\alpha)p_{d_0}$ becomes stable under the shared belief. Hence we show that for sufficiently large δ, there exists $\alpha^S \in (0,1)$ that makes $p_{c0d0}(\alpha^S) := a^S p_{c_0} + (1-\alpha^S)p_{d_0}$ a Nash equilibrium and moreover for some neighborhood U of α^S, for any $\alpha \in U$,

$$\alpha \gtreqless \alpha^S \iff v(d_0; p_{c0d0}(\alpha)) \gtreqless v(c_0; p_{c0d0}(\alpha)). \tag{7}$$

For any $\alpha \in (0,1)$,the average payoffs are as follows.

$$\begin{aligned} v(c_0; p_{c0d0}(\alpha)) &= \frac{\alpha c/(1-\delta^2) + (1-\alpha)\ell}{\alpha/(1-\delta^2) + 1 - \alpha} \\ v(d_0; p_{c0d0}(\alpha)) &= \alpha g + (1-\alpha) d. \end{aligned}$$

Hence the (relative) payoff difference becomes a quadratic function of α as follows. (See also Figure 2.)

$$\begin{aligned} F(\alpha) &:= \{v(c_0; p_{c0d0}(\alpha)) - v(d_0; p_{c0d0}(\alpha))\}\{\alpha/(1-\delta^2) + 1 - \alpha\}\{1-\delta^2\} \\ &= -\delta^2(g-d)\alpha^2 + \alpha\{c - \delta^2 d + (d-g-\ell)(1-\delta^2)\} - (1-\delta^2)(d-\ell). \end{aligned}$$

Clearly $F(0) < 0$ and $F(1) < 0$ but if the discriminant is positive, there exist two solutions in $(0,1)$. The larger solution satisfies (7). It can be shown by computation that the discriminant approaches to $(c-d)^2 > 0$ as δ approaches to 1. □

Proof of Proposition 5: Let $p_0^{\infty}(\alpha)$ be a strategy distribution consisting of infinitely many trust-building strategies $\{c_0, c_1, c_2, \ldots\}$ such that c_T-strategy has the share $\alpha(1-\alpha)^T$ in the distribution. Since c_T-strategies with $T \geqq 1$ behave the same way against c_0-strategy, i.e., they play D in the first period and then the partnership ends, it is obvious that for any $\alpha \in (0,1)$,

$$v(c_0; p_{cd0}(\alpha)) = v(c_0; p_0^1(\alpha)) = v(c_0; p_0^{\infty}(\alpha)).$$

Lemma 3. *For any* $\alpha \in (0,1)$*,*

$$v(c_0; p_0^{\infty}(\alpha)) = v(c_1; p_0^{\infty}(\alpha)) \iff v(c_0; p_0^{\infty}(\alpha)) = v(d_0; p_{cd0}(\alpha)).$$

Proof of Lemma 3: Recall the average payoff of c_0-strategy and d_0-strategy as functions of α (the fraction of c_0-strategy in the society) are as follows.

$$v(c_0; p_{cd0}(\alpha)) = v(c_0; p_0^{\infty}(\alpha)) = \frac{\alpha \frac{c}{1-\delta^2} + (1-\alpha)\ell}{\alpha \frac{1}{1-\delta^2} + 1 - \alpha}, \tag{8}$$

$$v(d_0; p_{cd0}(\alpha)) = \alpha g + (1-\alpha) d. \tag{9}$$

As for the average payoff of c_1-strategy under the infinite-strategy distribution, note that

- With probability α, c_1 is matched with c_0 and receive $V(c_1, c_0) = g$ while the partnership lasts $L(c_1, c_0) = 1$ period.

- With probability $1-\alpha$, c_1 is matched with either c_1 or c_T with $T \geqq 2$. For any $T \geqq 2$, c_T behaves the same way against c_1. In any case, in the first period c_1-strategy receives d. From the second period on, if the partner was c_1, then they enter the cooperation phase and receive $c/(1-\delta^2)$ for $L(c_1, c_1) = 1/(1-\delta^2)$ periods. If the partner was c_T with $T \geqq 2$, then c_1-strategy receives ℓ and ends the partnership.

Therefore the average payoff of c_1-strategy under the infinite-strategy distribution $p_0^\infty(\alpha)$ is

$$v(c_1; p_0^\infty(\alpha)) = \frac{\alpha g + (1-\alpha)[d + \delta^2\{\alpha \frac{c}{1-\delta^2} + (1-\alpha)\ell\}]}{\alpha + (1-\alpha)[1 + \delta^2\{\alpha \frac{1}{1-\delta^2} + (1-\alpha)\}]}. \tag{10}$$

For notational simplicity, let $L(c_0) = \alpha\frac{1}{1-\delta^2} + 1 - \alpha$ and $V(c_0) = \alpha\frac{c}{1-\delta^2} + (1-\alpha)\ell$. Then, using (8), (9), and (10), we have that

$$\begin{aligned}
& v(c_0; p_0^\infty(\alpha)) = v(c_1; p_0^\infty(\alpha)) \\
\iff \quad & \frac{V(c_0)}{L(c_0)} = \frac{\alpha g + (1-\alpha)\delta^2 V(c_0)}{1 + (1-\alpha)\delta^2 L(c_0)} \\
\iff \quad & \{1 + (1-\alpha)\delta^2 L(c_0)\}V(c_0) = \{\alpha g + (1-\alpha)d + (1-\alpha)\delta^2 V(c_0)\}L(c_0) \\
\iff \quad & V(c_0) = \{\alpha g + (1-\alpha)d\}L(c_0) \\
\iff \quad & v(c_0; p_0^\infty(\alpha)) = v(d_0; p_{cd0}(\alpha)).
\end{aligned}$$

■

This Lemma implies that the equilibrium payoff of c_0-strategy under the bimorphic distribution $p_{cd}(\alpha^*)$ is the same as the equilibrium payoff of c_0-strategy under $p_0^\infty(\alpha^*)$.

Proposition 4 of Greve-Okuno (2009) shows that

$$v(c0; p_0^1(\alpha_0^1)) < v(c_0; p_0^\infty(\alpha^*)).$$

Hence

$$v(c0; p_0^1(\alpha_0^1)) < v(c_0; p_0^\infty(\alpha^*)) = v(c_0; p_{cd}(\alpha^*)).$$

□

REFERENCES

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. and Okuno-Fujiwara, M. (2009), "Voluntarily Seperable Repeated Prisoner's Dilemma", forthcoming in *Review of Economic Studies.* (Available at http://www.restud.com/uploads/papers/MS11308-3%20manuscript.pdf)

Ghosh, P. and Ray D. (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.

Kandori, M., Mailath, G. and Rob, R. (1993), "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica* , **61**, 29-56.

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

Kreps, D. and Wilson, R. (1982), "Sequential Equilibria", *Econometrica*, 50: 863-894.

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

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

Myerson, R. (1978), "Refinements of the Nash Equilibrium Concept", *International Journal of Game Theory*, **7**, 73-80.

Robson, A. (1990), "Efficiency in Evolutionary Games: Darwin, Nash and the Secret Handshake," *Journal of Theoretical Biology*, **144**, 379-396.