CIRJE-F-415

Voluntarily Separable Prisoner's Dilemma

Takako Fujiwara-Greve
Keio University
and
Norwegian School of Management BI

Masahiro Okuno-Fujiwara
University of Tokyo

April 2006

Voluntarily Separable Prisoner's Dilemma*

by

Takako Fujiwara-Greve
Dept. of Economics, Keio University
2-15-45 Mita, Minatoku, Tokyo 108-8345 JAPAN
and
Dept. of Economics, Norwegian School of Management BI
N-0442 Oslo, NORWAY
takakofg@econ.keio.ac.jp

and

Masahiro Okuno-Fujiwara
Dept. of Economics, University of Tokyo
7-3-1 Hongo, Bunkyoku, Tokyo 113-0033 JAPAN
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. There is no information flow across matches, and players are boundedly rational. We extend the notion of Neutrally Stable Distribution (NSD) to fit for our model. When the component game is a prisoner's dilemma, NSD requires some *trust-building periods* to defect at the beginning of a partnership. We find that *polymorphic* 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 polymorphic NSD is higher.

Key words: voluntary separation, prisoner's dilemma, evolution, trust.
JEL classification number: C 73

*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, 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.

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. 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 consider boundedly rational players who are endowed with a pure strategy and analyze evolutionary stability of strategy distributions. Since our model is an extensive form game, there are many strategies that only differ in the off-path decision nodes. Hence invasion concept needs to be carefully defined. We extend Neutrally Stable Distribution (NSD) concept, 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 (*polymorphic*) 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 polymorphic NSDs. We found that polymorphic 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 polymorphic NSDs are more efficient than the most efficient monomorphic NSD.

The existence of polymorphic NSDs in a homogeneous population provides an evolutionary foundation to incomplete information models of voluntarily separable repeated games (e.g., Ghosh and Ray, 1997, and Rob and Yang, 2005). Diverse strategies co-exist, discipline each other, and the shortest trust-building strategy can survive thanks to efficient outcome when meeting the same-type partner.

Extensions include cheap-talk model and other types of strategy distributions and component games. The trust-building periods can be viewed as a signal to distinguish cooperative strategies from others. Then it is natural to extend the model to allow cheap-talk. When cheap-talk is introduced at the beginning of a new partnership, the most efficient symmetric NSD is the unique symmetric NSD that cannot be invaded by equilibrium entrants (Swinkels, 1992). We also mention how coordinated action profiles over time (such as alternating (C, D) and (D, C)) can be sustained. This leads to the analysis of asymmetric stage games such as Hawk-Dove.

This paper is organized as follows. In Section 2, we introduce the formal model and stability concepts. In Section 3, we identify the shortest trust-building periods for monomorphic NSDs. In Section 4, we identify shortest trust-building periods for various polymorphic NSDs. In Section 5 we discuss extensions including the cheap-talk model and give concluding remarks.

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.

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

[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

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[2] $g > c > d > \ell$ 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.[3] Let $H_t := \{C, D\}^{2(t-1)}$ be *the set of partnership histories at the beginning of* $t \geq 2$ and let $H_1 := \{\emptyset\}$.

DEFINITION. A *pure strategy* s of VSPD specifies $(x_t, y_t)_{t=1}^{\infty}$ where:
$x_t : H_t \to \{C, D\}$ specifies an action choice $x_t(h_t) \in \{C, D\}$ given the partnership history $h_t \in H_t$, and
$y_t : H_t \times \{C, D\}^2 \to \{k, e\}$ specifies whether to keep or end the partnership, depending upon the partnership history $h_t \in H_t$ and the current period action profile.

The 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

[2]We make a remark on the case of $2c < g + \ell$ in the concluding remark.

[3]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.

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". Moreover, by looking at the strategy distributions in the matching pool, we can directly compute life-time payoffs of players easily.

2.2. *Life-time and Average Payoff in a Match*

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, 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,

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

2.3. *Life-time and Average Payoff in the Matching Pool*

Next we show the structure of the life-time 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

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

where $L = 1 + \delta + \delta^2 + \cdots = \frac{1}{1-\delta}$ is the number of total days s expects to live.

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[4] as a convex combination of "in-match" 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 $v(s;p) = v^I(s,s')$.

2.4. *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 $s' \in \boldsymbol{S}$,

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

[4]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.

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 $\tilde{d}$ which plays D at $t = 1$ and ends the partnership for any observation at $t = 1$. For $t \geq 2$, which is off-path, specify arbitrary actions. Then any $\tilde{d}$-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 a $\tilde{d}$-strategy is a Nash equilibrium. Against a $\tilde{d}$-strategy, any strategy must play one-shot Prisoner's Dilemma. Hence, any strategy that starts with C in $t = 1$ earns strictly lower average payoff than that of a $\tilde{d}$-strategy, and any strategy that starts with D in $t = 1$ earns the same average payoff as that of a $\tilde{d}$-strategy.

2.5. *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 show that any $\tilde{d}$-strategy is not NSS, even though it constitutes a symmetric Nash equilibrium. Hence NSD concept selects among Nash equilibria in our model.

LEMMA 2. *Any 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.

2.6. *Simple, Trust-building Strategies*

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 played by boundedly rational players in an 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).

Intuitively, we focus on generalized versions of c_1-strategy that can invade the population of a myopic $\tilde{d}$-strategy. To formalize, we first define *simple strategies*, which have a set of acceptable paths and end a partnership as soon as a deviation from acceptable paths is observed. Simple strategies, however, do not restrict the equilibrium outcomes because the equilibrium continuation payoffs cannot be lower than the continuation payoff after voluntary separation.

Let $\Omega = \cup_{t=1}^{\infty}(\{C, D\} \times \{C, D, \emptyset\})^{(t-1)}$. Interpret that the first coordinate is the player's own action and the second coordinate is the current partner's "acceptable" action. The $\emptyset$ means that any action by the partner is not acceptable, i.e., the strategy intends to end the partnership regardless of the observation at that point. For any $q \in \Omega$, let $|q|$ be the length of the sequence q, i.e., the number of action profiles contained in q.

DEFINITION. $Q \subset \Omega$ is the set of *acceptable paths* if,

(1) for any $q, q' \in Q$ and any $t = 1, 2, \ldots, \min\{|q|, |q'|\}$, if $(q(1), \ldots, q(t-1)) = (q'(1), \ldots, q'(t-1))$, then $q_1(t) = q'_1(t)$;

(2) for any $q \in Q$, if $q_2(t) = \emptyset$ for some t, then $|q| = t$.

The first condition guarantees that the action is uniquely determined after any acceptable observed path. The second condition means that if a strategy intends to end the partnership at t, then the specification of the acceptable path ends there.

DEFINITION. For any set of acceptable paths $Q \subset \Omega$, a strategy $s(Q) \in \boldsymbol{S}$ is a *simple strategy* if, in each period t,
(a) in the stage game, it plays according to the unique $q_1(t)$ generated by Q and the observed path; and
(b) in the continuation decision phase, it keeps the partnership if and only if the observed path is the same as the first t components of some $q \in Q$.

An extension of the ordinary C-trigger strategy to our model is a simple strategy with a singleton set of acceptable path

$$Q_{tr} = \{((C, C), (C, C), \ldots)\}.$$

Any myopic $\tilde{d}$-strategy is also a simple strategy with $Q_{\tilde{d}} = \{((D, \emptyset))\}$.

Next, we define trust-building strategies, a generalization of c_1-strategy.

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 simple strategy with the singleton set of acceptable path

$$Q_{c_T} = \{(\overbrace{(D,D),\ldots,(D,D)}^{\text{T times}},(C,C),(C,C),\ldots)\}.$$

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 simple 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.

3. MONOMORPHIC STRATEGY DISTRIBUTIONS

We first consider *monomorphic* strategy distributions, consisting of a single c_T-strategy. The literature of voluntarily separable repeated games has focused on similar symmetric strategy distributions.

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 deviations are 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.

Suppose that one plays D at some point during the cooperation periods. The player receives g but returns to the matching pool if he does not die. The continuation payoff is thus $g + \delta V(c_T; p_T)$. By contrast, the expected continuation payoff of c_T-strategy during the cooperation periods is $L(c_T, c_T)c + \delta(1-\delta)(1 + \delta^2 + \cdots)V(c_T; p_T) = L(c_T, c_T)c + \delta(1-\delta)L(c_T, c_T)V(c_T; p_T)$. Therefore, one-step deviation during cooperation periods is not better than c_T-strategy if and only if

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

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.

Now we prove that in fact the Best Reply Condition (5) is the only condition that is required for p_T to be a Nash equilibrium. Let *on-path history* at a decision node of $t = 1, 2, 3, \ldots$, be the play path until the decision node of the t-th period in a match of two c_T-strategies. That is, the on-path history in PD in periods $t \leq T$ is $(D, D)^{t-1}$ and in periods $t \geq T+1$ is $\{(D, D)^T, (C, C)^{(t-T-1)}\}$. The on-path history at the continuation decision phase is similarly defined.

LEMMA 3. *Take an arbitrary $T = 1, 2, 3, \ldots$. Let p_T be the stationary strategy distribution in the matching pool, consisting only of c_T-strategy.*

(a) Any strategy that ends the match in some period $t = 1, 2, \ldots$ along on-path history is not a best reply against p_T.

(b) *Any strategy that chooses* C *at some* $t < T + 1$ *along on-path history is not a best reply against* p_T*.*

(c) *Let* s *be any strategy that chooses* D *at some* $t \geq T + 1$ *along on-path history. Then* $v(c_T; p_T) \geq v(s; p_T)$ *if and only if* $v(c_T; p_T) \leq v^{BR}$*.*

PROOF: See Appendix.

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 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 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: (Can be omitted.) Lemma 3 implies that 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 (5) holds, i.e., $T \geq \underline{\tau}(\delta)$. 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.

LEMMA 4. *Suppose $p \in \mathcal{P}(\boldsymbol{S})$ is a Nash equilibrium. If a pure strategy $s' \in S$ invades p for some $\epsilon > 0$, then s' is an alternatvie best reply to p, i.e., $s' \in BR(p)$.*

PROOF: (Obvious from (1). Can be omitted.) See Appendix.

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, Lemma 3 (c) implies that such strategies are not alternative best reply. Therefore c_T-strategy is NSS for this case.

When $\underline{\tau}(\delta)$ is an integer, the Nash equilibrium $p_{\underline{\tau}(\delta)}$ has alternative best replies, among which $c_{\underline{\tau}(\delta)+1}$ earns the highest payoff when meeting itself. It suffices to check if $c_{\underline{\tau}(\delta)+1}$-strategy cannot invade $p_{\underline{\tau}(\delta)}$.

Below we first show general properties of the average values of c_T-strategy and c_{T+1}-strategy for any T when both of these are present in the matching pool. This is useful in the later analysis as well. After that we show a condition that $c_{\underline{\tau}(\delta)}$-strategy earns a higher payoff when the fraction of $c_{\underline{\tau}(\delta)+1}$-strategy is sufficiently small.

For any T, let $p_T^{T+1}(\alpha) = \alpha p_T + (1-\alpha)p_{T+1}$ be a two-strategy distribution of c_T and c_{T+1}.

LEMMA 5. *For any $\delta \in (\underline{\delta}(\infty), 1)$ and any $T = 0, 1, 2 \ldots$, $v(c_T; p_T^{T+1}(\alpha))$ is strictly increasing and concave function of α.*

PROOF: (By differentiation. Can be omitted.) See Appendix.

LEMMA 6. *For any $\delta \in (\underline{\delta}(\infty), 1)$ and any $T = 0, 1, 2 \ldots$ such that $T \leq \underline{\tau}(\delta)$, $v(c_{T+1}; p_T^{T+1}(\alpha))$ is strictly increasing and convex function of α.*

PROOF: (By differentiation. Can be omitted.) See Appendix.

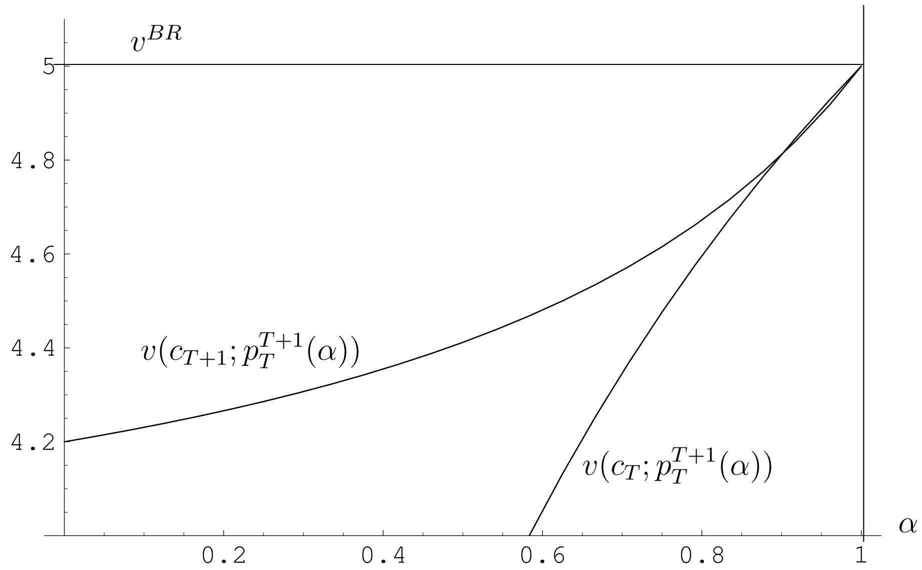

FIGURE 1. – The value functions of c_T-strategy and c_{T+1}-strategy when $T = \underline{\tau}(\delta)$. (Parameter values: $g = 10, c = 6, d = 1, \ell = -1, \delta = \frac{2}{\sqrt{5}}, T = \underline{\tau}(\delta) = 1.$)

The intuition of the concavity and convexity of the average payoffs of c_T and c_{T+1}-strategy respectively is as follows. As α decreases from 1 towards 0, c_T-strategy gets exploited by c_{T+1}-strategy more often and the exploitation accelerates as the fraction of c_{T+1}-strategy increases. Hence the average value of c_T-strategy drops more as α decreases. On the other hand, as α increases from 0 to 1, c_{T+1}-strategy benefits more and more by the increased probability of meeting c_T-strategy. Thus the average value of c_{T+1}-strategy increases more as α increases.

Thanks to the concavity and convexity, c_{T+1}-strategy cannot invade p_T if and only if the slope of $v(c_T; p_T^{T+1}(\alpha))$ is strictly smaller than the slope of $v(c_{T+1}; p_T^{T+1}(\alpha))$ at $\alpha = 1$, see Figure 1.

LEMMA 7. *Take any* $\delta \in (\underline{\delta}(\infty), 1)$. *Let* $T = \underline{\tau}(\delta)$. *Then*

$$\frac{\partial v(c_T; p_T^{T+1}(\alpha))}{\partial \alpha}\Big|_{\alpha=1} < \frac{\partial v(c_{T+1}; p_T^{T+1}(\alpha))}{\partial \alpha}\Big|_{\alpha=1}$$

if and only if

$$[1 - \delta^{2(T+1)}](g - \ell) < c - d. \tag{6}$$

Proof: (By computation. Can be omitted.) See Appendix.

Hence, if we 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 $\underline{\tau}(\delta)<\hat{\tau}(\delta)$.

To interpret (6), notice that $L(c_T,c_T)=1+\delta^2+\cdots$ and $L(c_{T+1},c_T)=1+\delta^2+\cdots+\delta^{2T}$. Hence the condition (6) is equivalent to

$$(g-\ell)L(c_{T+1},c_T)<(c-d)L(c_T,c_T) \tag{7}$$

at $T=\underline{\tau}(\delta)$. The RHS of (7) can be interpreted as the relative merit of c_T-strategy against c_{T+1}-strategy (to start cooperating one period early when meeting itself) and the LHS is the relative merit of c_{T+1}-strategy when meeting c_T-strategy.

As δ increases (when G is fixed), T must increase to keep the equality (6). Thus $\hat{\tau}$ is an increasing function of δ and goes to ∞ as $\delta\to 1$.

Lemma 8. *There exists a unique $\delta^*\in(\underline{\delta}(\infty),1)$ that satisfies*

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

Proof: (Can be omitted. See Figure 2.) Recall that $\underline{\tau}(\delta)$ is decreasing in δ. When $\delta=1$, $\hat{\tau}(1)=\infty>\underline{\tau}(1)$. The δ satisfying $\hat{\tau}(\delta)=0$ is $\delta=\sqrt{\frac{g-c+d-\ell}{g-\ell}}>\sqrt{\frac{g-c}{g-d}}=\underline{\delta}(\infty)$, where the inequality obtains by computation. Hence when δ is close to $\underline{\delta}(\infty)$, $\hat{\tau}(\delta)<\underline{\tau}(\delta)$. Q.E.D.

In summary, most of the symmetric Nash equilibirum strategies are NSS except at some boundary values.

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 $T>\underline{\tau}(\delta)$.

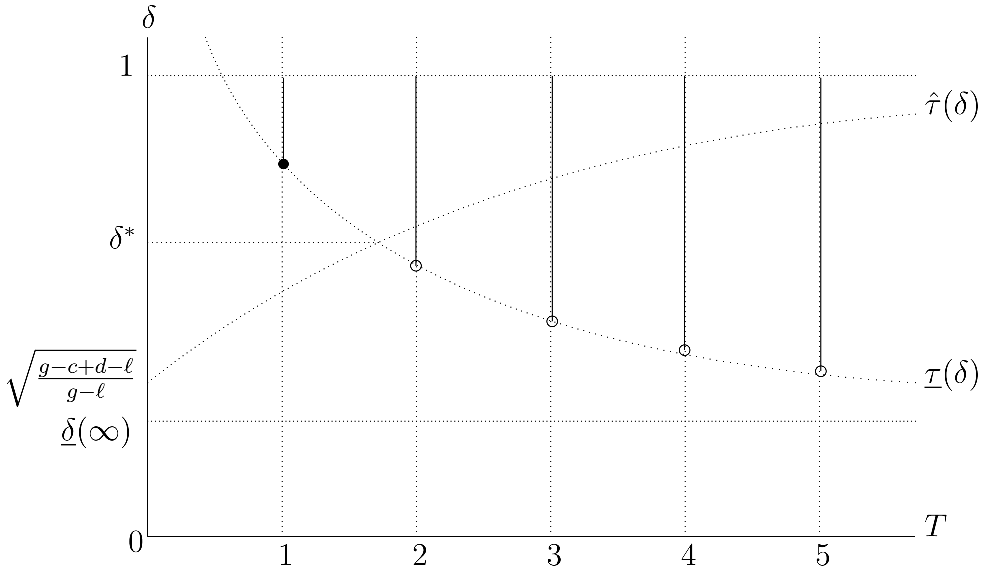

FIGURE 2. – Parametric summary of monomorphic NSS

4. POLYMORPHIC STRATEGY DISTRIBUTIONS

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. (See concluding remark Section 5.5.) 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. *Bimorphic Distribution*

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. Therefore we need:

Payoff Equalization: there exists $\alpha_T^{T+1} \in (0,1)$ and a neighborhood U of α_T^{T+1} such that for any $\alpha \in U$

$$\alpha \gtreqless \alpha_T^{T+1} \iff v(c_{T+1}; p_T^{T+1}(\alpha)) \lesseqgtr v(c_T; p_T^{T+1}(\alpha)). \tag{8}$$

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 or c_{T+1}-strategy to enter cooperation periods (i.e., play C at least once at $T+1$ or $T+2$) and then play D to earn g for sure. 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) holds.

LEMMA 9. *Best Reply Condition: Any one-step deviation strategy from c_T or c_{T+1}-strategy does not earn higher average payoff than c_T or c_{T+1} if and only if*

$$v(c_T; p_T^{T+1}(\alpha_T^{T+1})) \leq v^{BR}. \tag{9}$$

PROOF: (By computation. Can be omitted.) See Appendix.

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 3. Clearly, there is no bimorphic NSD with the support $\{c_{\underline{\tau}(\delta)}, c_{\underline{\tau}(\delta)+1}\}$. For T *slightly* below $\underline{\tau}(\delta)$, 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 $T < \hat{\tau}(\delta)$, using a similar logic to Lemma 7.

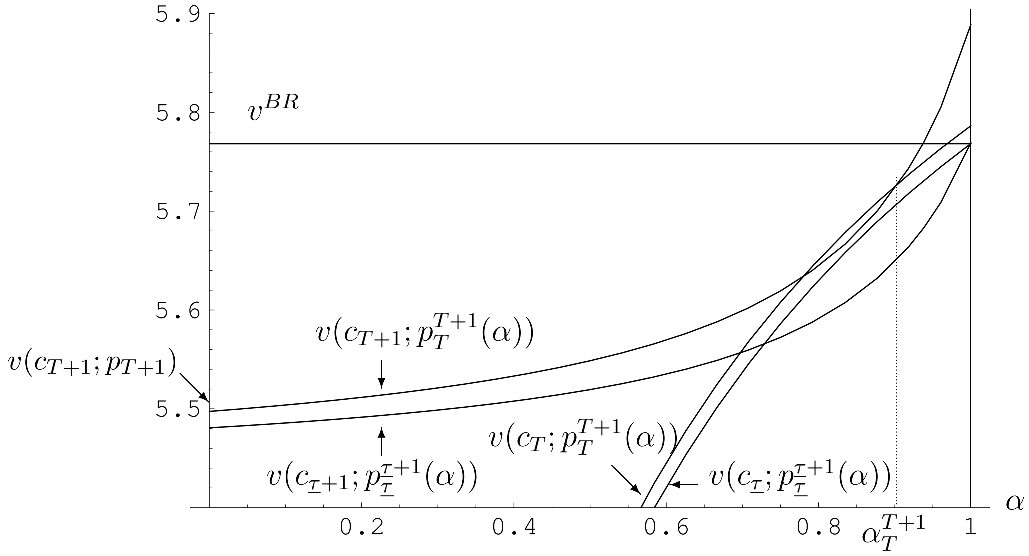

FIGURE 3. – The existence of a bimorphic NSD as T is slightly below $\underline{\tau}(\delta)$. (Parameter values: $g = 10, c = 6.1, d = 2.1, \ell = 2, \delta = 0.96, T = 1, \underline{\tau}(\delta) \approx 1.06$.)

In a bimorphic NSD, the shortest trust-building periods is shorter than any of monomorphic NSD because earlier opportunity to deviate is offset by the possible exploitation by c_{T+1}-strategy in the future match.

PROPOSITION 3. *For any $\delta > \delta^*$, there exists $\tau_2(\delta)$ such that $\tau_2(\delta) < \underline{\tau}(\delta)$, and, for any T such that $\tau_2(\delta) < T < \underline{\tau}(\delta)$, 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 $\tau_2(\delta)$ warrants that the payoff-equalizing $\alpha_T^{T+1}(\delta)$ exists. Then one can prove 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 $(\tau_2(\delta), \underline{\tau}(\delta))$ contains an integer. Figure 3 is a numerical example of such G.

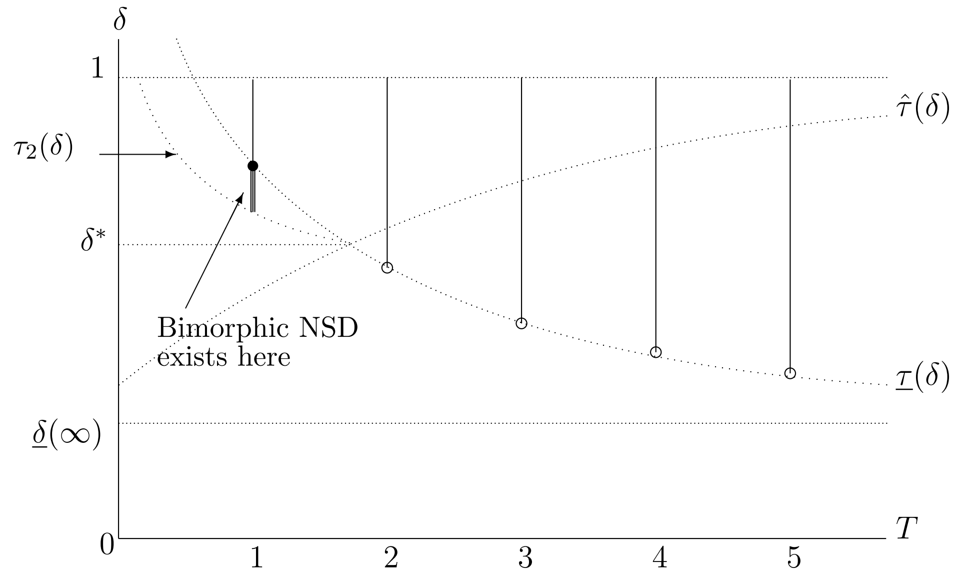

FIGURE 4. – Bimorphic NSD.

4.2. *Higher Efficiency of Bimorphic NSD*

For a given $\delta > \delta^*$, the shortest trust-building periods in the support of a bimorphic NSD, if it exists, 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 $p_T^{T+1}(\alpha_T^{T+1}(\delta))$. The average payoff of c_{T+1} strategy as a NSS is

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

Lemma 6 shows that $v(c_{T+1}; p_T^{T+1}(\alpha))$ is an increasing function of α, and thus

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

since $\alpha_T^{T+1}(\delta) > 0$. (See Figure 3.) Hence bimorphic NSDs, if they exist, are more efficient than any monomorphic NSS, thanks to earlier cooperation, even though equilibrium break-up occurs.

4.3. *Staggered Distribution: Finite Support*

We can extend the analysis of the bimorphic NSDs for NSDs with a finite support of the form $\{c_T, c_{T+1}, \ldots, c_{T+K}\}$, which we call a $(K+1)$-morphic distribution. First, we derive *trimorphic* NSDs with the support of $\{c_T, c_{T+1}, c_{T+2}\}$ as a benchmark.

Let $p_T^{T+2}(\alpha,\beta) = \alpha p_T + (1-\alpha)\beta p_{T+1} + (1-\alpha)(1-\beta)p_{T+2}$ be a trimorphic distribution. The Payoff Equalization condition should be derived backwards. Given the fraction α of c_T-strategy, find $\beta^*(\alpha) \in (0,1)$ such that there exists a neighborhood U of $\beta^*(\alpha)$ and for any $\beta \in U$,

$$\beta \gtreqless \beta^*(\alpha) \iff v(c_{T+2}; p_T^{T+2}(\alpha,\beta)) \gtreqless v(c_{T+1}; p_T^{T+2}(\alpha,\beta)). \tag{10}$$

The $\beta^*(\alpha)$ exists if and only if the quadratic equation of β,

$$\begin{aligned}
&\{v(c_{T+1}; p_T^{T+2}(\alpha,\beta)) - v(c_{T+2}; p_T^{T+2}(\alpha,\beta))\}\\
&\times\{\alpha L(c_{T+1}, c_T) + (1-\alpha)\beta L(c_{T+1}, c_{T+1}) + (1-\alpha)(1-\beta)L(c_{T+1}, c_{T+2})\}\\
&\times\{\alpha L(c_{T+2}, c_T) + (1-\alpha)\beta L(c_{T+2}, c_{T+1}) + (1-\alpha)(1-\beta)L(c_{T+2}, c_{T+2})\} = 0
\end{aligned}$$

has two solutions within $(0,1)$. The larger one is $\beta^*(\alpha)$.

Then we find $\alpha_T^{T+2} \in (0,1)$ (dependent on δ) and its neighborhood W such that for any $\alpha \in W$,

$$\alpha \gtreqless \alpha_T^{T+2}(\delta) \iff v(c_{T+1}; p_T^{T+2}(\alpha, \beta^*(\alpha))) \gtreqless v(c_T; p_T^{T+2}(\alpha)). \tag{11}$$

Note that the average payoff of c_T-strategy only depends on α, since c_{T+1} and c_{T+2} behave the same way against c_T. The payoff-equalizing $\alpha_T^{T+2}(\delta)$ exists if and only if the intersection exists between $v(c_T; p_T^{T+2}(\alpha))$ and $v(c_{T+1}; p_T^{T+2}(\alpha, \beta^*(\alpha)))$, both of which are functions of α only. See Figure 5.

The *Best Reply Condition* is derived in the same way as before. Any c_{T+k}-strategy $(k = 0, 1, 2)$ is optimal if and only if

$$\begin{aligned}
& g + \delta V(c_{T+k}; p_T^{T+2}(\alpha, \beta^*(\alpha))) \le \frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2} V(c_{T+k}; p_T^{T+2}(\alpha, \beta^*(\alpha)))\\
\iff \quad & v(c_{T+k}; p_T^{T+2}(\alpha, \beta^*(\alpha))) \le \frac{c - (1-\delta^2)g}{\delta^2} = v^{BR}.
\end{aligned} \tag{12}$$

It can be shown that the average payoff of c_{T+1} under a payoff-equalizing trimorphic distribution (i.e., under $(\alpha, \beta^*(\alpha))$) intersects with v^{BR} at α where it intersects with v^{BR} under a bimorphic distribution. Moreover, since there are exploiters for c_{T+1}-strategy (namely c_{T+2}-strategy), the average value of c_{T+1}-strategy is lower

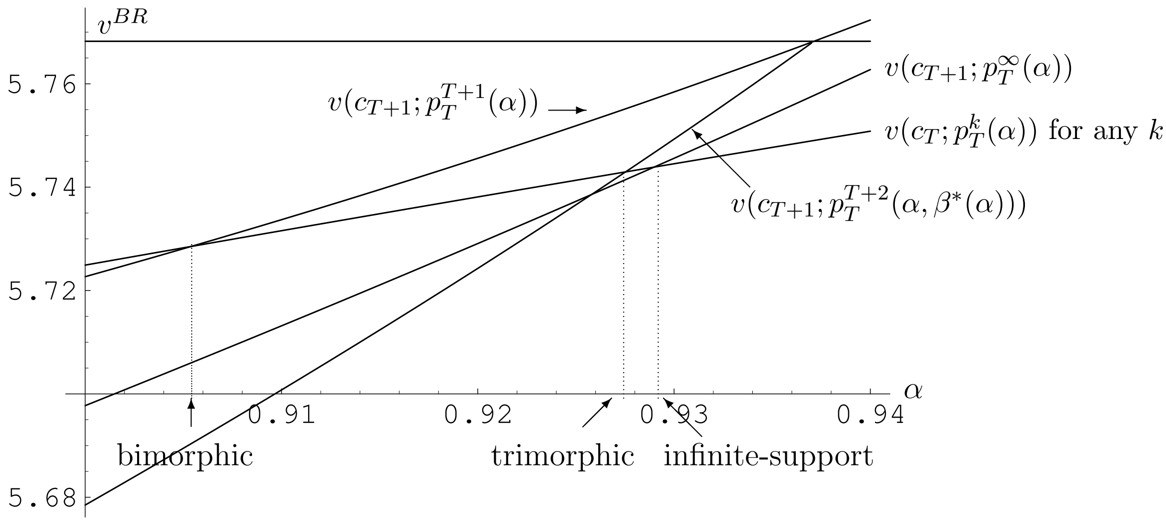

FIGURE 5. – Existence of Polymorphic NSDs.

(Parameter values: $g = 10, c = 6.1, d = 2.1, \ell = 2, \delta = 0.96, T = 1, \underline{\tau}(\delta) \approx 1.06$.)

under the trimorphic distribution than under the bimorphic distribution with the same α. See Figure 5.

LEMMA 10. *For any $\delta > \delta^*$ and any $T < \underline{\tau}(\delta)$, let $\alpha^*_{T+1}(v^{BR})$ be the fraction of c_T-strategy that solves*

$$v(c_{T+1}; p_T^{T+1}(\alpha)) = v^{BR}.$$

Then, $\beta^*(\alpha^*_{T+1}(v^{BR})) = 1$ and for any $\alpha < \alpha^*_{T+1}(v^{BR})$,

$$v(c_{T+1}; p_T^{T+2}(\alpha, \beta^*(\alpha))) < v(c_{T+1}; p_T^{T+1}(\alpha)).$$

PROOF: See Appendix.

Hence if a bimorphic NSD p_T^{T+1} exists, then a trimorphic NSD exists. (But not vice versa.)

PROPOSITION 4. *For any $\delta > \delta^*$ there exists $\tau_3(\delta) < \tau_2(\delta)$ and, for any T such that $\tau_3(\delta) < T < \underline{\tau}(\delta)$, there exists a trimorphic NSD of the form $p_T^{T+2}(\alpha_T^{T+2}, \beta^*(\alpha_T^{T+2}))$, where $\alpha_T^{T+2} \in (0, 1)$.*

PROOF: Lemma 10 implies that if there exists α such that $v(c_{T+1}; p_T^{T+1}(\alpha)) = v(c_T; p_T^{T+1}(\alpha)) < v^{BR}$ which satisfies (8), then there exists α such that $v(c_{T+1}; p_T^{T+2}(\alpha, \beta^*(\alpha))) = v(c_T; p_T^{T+2}(\alpha)) < v^{BR}$ which satisfies (11). Moreover, even if $T = \tau_2(\delta)$ so that there is no bimorphic NSD, the payoff-equalizing α_T^{T+2} exists to warrant a trimorphic NSD. Therefore the lower bound to T for the existence of a trimorphic NSD is lower than $\tau_2(\delta)$. Q.E.D.

In addition, the average payoff of the trimorphic NSD is greater than that of the bimorphic NSD, thanks to the increasing nature of $v(c_T; p_T^{T+1}(\alpha))$ in α. This means that even if the shortest trust-building periods is the same between a bimorphic NSD and a trimorphic NSD, more diversity gives higher efficiency. The intuition is that with more variety of strategies (i.e., more exploiters) in the society, the strategy with the shortest trust-building periods must increase its fraction in equilibrium.

In general, let

$$p_T^{T+K}(\alpha, \beta_1, \beta_2, \ldots, \beta_K) = \alpha p_T + (1-\alpha) \sum_{k=1}^{K} \Pi_{m=1}^{k-1}(1-\beta_m)\beta_k p_{T+k}$$

be a $(K+1)$-morphic distribution. Define the Payoff-Equalizing β_k^*'s $(k = 1, 2, \ldots, K)$ as follows. For notational simplicity, let us write a vector $\beta_1^k = (\beta_1, \ldots, \beta_k)$ for any $k = 1, 2, \ldots, K$.

Given the fractions (α, β_1^{K-1}), define $\beta_K^*(\alpha, \beta_1^{K-1}) \in (0, 1)$ that makes c_{T+K} and c_{T+K-1} equivalent and un-invadable by strategies with the same on-path actions:

$$\begin{aligned} & \beta_K \gtreqless \beta_K^*(\alpha, \beta_1^{K-1}) \\ \iff \quad & v(c_{T+K-1}; p_T^{T+K}(\alpha, \beta_1^{K-1}, \beta_K)) \gtreqless v(c_{T+K}; p_T^{T+K}(\alpha, \beta_1^{K-1}, \beta_K)), \end{aligned} \tag{13}$$

for any β_K in some neighborhood of $\beta_K^*(\alpha, \beta_1^{K-1})$. Similarly, given (α, β_1^{K-2}) and $\beta_K^*(\cdot)$, define $\beta_{K-1}^*(\alpha, \beta_1^{K-2}) \in (0, 1)$:

$$\begin{aligned} & \beta_{K-1} \gtreqless \beta_{K-1}^*(\alpha, \beta_1^{K-2}) \\ \iff \quad & v(c_{T+K-2}; p_T^{T+K}(\alpha, \beta_1^{K-2}, \beta_{K-1}, \beta_K^*(\alpha, \beta_1^{K-2}, \beta_{K-1})) \\ & \gtreqless v(c_{T+K-1}; p_T^{T+K}(\alpha, \beta_1^{K-2}, \beta_{K-1}, \beta_K^*(\alpha, \beta_1^{K-2}, \beta_{K-1}))), \end{aligned}$$

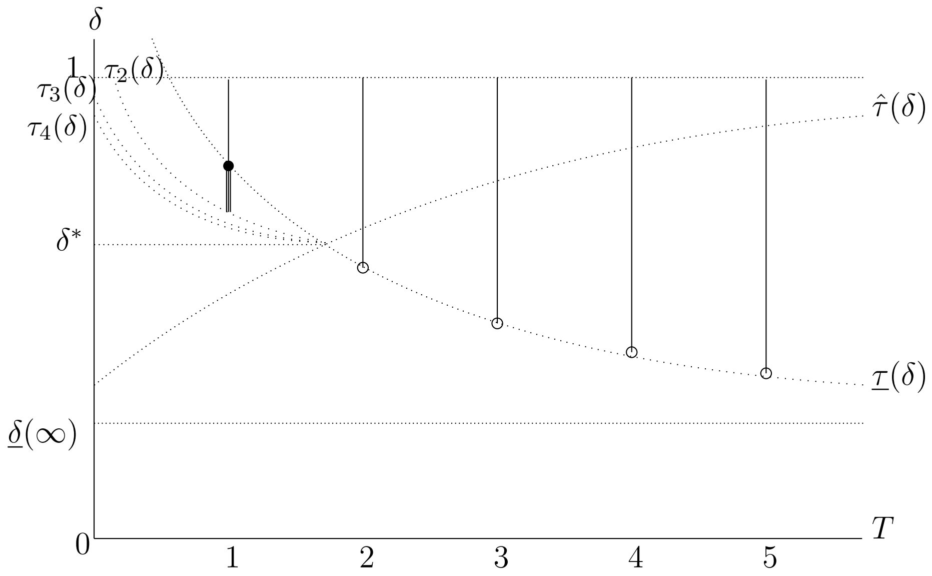

FIGURE 6. – Finite Support Polymorphic NSDs.

for any β_{K-1} in some neighborhood of $\beta^*_{K-1}(\alpha, \beta_1^{K-2})$. After computing $\beta^*_k(\alpha, \beta_1^{k-1})$ for all $k = 1, 2, \ldots, K$, finally find $\alpha_T^{T+K} \in (0, 1)$ (dependent on δ) such that

$$\alpha \gtreqless \alpha_T^{T+K}(\delta) \iff v(c_{T+1}; p_T^{T+K}(\alpha, \beta_1^*(\alpha))) \gtreqless v(c_T; p_T^{T+K}(\alpha)), \tag{14}$$

for any α in some neighborhood of $\alpha_T^{T+K}(\delta)$. Note that the average payoff of c_T-strategy only depends on α, since $c_{T+1}, \ldots, c_{T+K}$ behave the same way against c_T, and the average payoff of c_{T+1}-strategy only depends on α and β_1.

The payoff-equalizing $\alpha_T^{T+K}(\delta)$ exists if and only if the intersection exists between $v(c_T; p_T^{T+K}(\alpha))$ and $v(c_{T+1}; p_T^{T+K}(\alpha, \beta_1^*(\alpha)))$. The *Best Reply Condition* is derived in the same way as before. Using the same logic as Lemma 10, a $(K+1)$-morphic NSD exists if K-morphic NSD exists. See Figure 6. In particular, for $T \in (\tau_2(\delta), \underline{\tau}(\delta))$, any K-morphic NSD exists for $K = 2, 3, \ldots$.

4.4. *Staggered Distribution: Infinite Support*

Finally we consider simple strategy distributions with the support $\{c_T, c_{T+1}, \ldots\}$, i.e., infinitely many variety of trust-building periods. We first prove that if a strategy distribution with the support $\{c_T, c_{T+1}, \ldots\}$ is to become a NSD, then the population distribution of c_t-strategies must be "geometric".

LEMMA 11. *For any $T < \infty$, let p be a stationary strategy distribution with the support $\{c_T, c_{T+1}, \ldots\}$. If $v(c_T; p) = v(c_{T+k}; p)$ for all $k = 1, 2, \ldots$, then there exists $\alpha \in (0, 1)$ such that the fraction of c_{T+k}-strategy is of the form $\alpha(1-\alpha)^k$ for each $k = 0, 1, 2, \ldots$.*

PROOF: See Appendix.

Denote the geometric distribution of $\{c_T, c_{T+1}, \ldots\}$ as $p_T^\infty(\alpha)$. If $p_T^\infty(\alpha)$ is the stationary strategy distribution in the matching pool and if c_T and c_{T+1} have the same average payoff, then all other strategies in the support have also the same payoff. The intuition is as follows. From the second period on, c_{T+1}-strategy behaves the same way as c_T-strategy against itself (there are T remaining periods of trust-building) and against longer trust-building strategies (it ends the partnership after T periods). The conditional probabilities of meeting itself and longer trust-building strategies are also the same as those of c_T-strategy.

Similarly, from the second period on, c_{T+2}-strategy behaves the same way as c_{T+1}-strategy against itself and against longer trust-building strategies. Therefore, if c_T and c_{T+1}-strategy have the same average payoff, all others have the same average payoff as well. (See Table II in the Appendix.)

LEMMA 12. *For any $T < \infty$ and any $\alpha \in (0, 1)$, if $v(c_T; p_T^\infty(\alpha)) = v(c_{T+1}; p_T^\infty(\alpha))$, then $v(c_{T+k}; p_T^\infty(\alpha)) = v(c_T; p_T^\infty(\alpha))$ for all $k = 1, 2, \ldots$.*

PROOF: (Can be omitted.) See Appendix.

It is straightforward to show that $v(c_{T+1}; p_T^\infty(\alpha)) < v(c_{T+1}; p_T^{T+1}(\alpha))$ for any $\alpha < \alpha_{T+1}(v^{BR})$ so that the payoff-equalizing α exists if a bimorphic NSD exists. The Best Reply Condition is derived as follows. Notice that for any period after T, playing D (after repeating (D, D)) is an on-path action. Hence the meaningful deviation strategies are those that play D after the cooperation periods started (that is, play D and keep the partnership if and only if (D, D) is observed for first $T + k$ periods, play C at least once, and then play D.) Using the continuation values at

$T+k+2$, such one-step deviation during the cooperation periods is not better than c_{T+k}-strategy if and only if

$$\begin{aligned} & g+\delta V(c_{T+k};p_T^\infty(\alpha)) \leq \frac{c}{1-\delta^2}+\frac{\delta(1-\delta)}{1-\delta^2}V(c_{T+k};p_T^\infty(\alpha)) \\ \iff \quad & v(c_{T+k};p_T^\infty(\alpha)) \leq \frac{c-(1-\delta^2)g}{\delta^2}=v^{BR}. \end{aligned} \tag{15}$$

PROPOSITION 5. *For any $\delta > \delta^*$ there exists $\tau^*(\delta) < \tau_2(\delta)$ and, for any T such that $\tau^*(\delta) < T < \underline{\tau}(\delta)$, there is a NSD of the form $p_T^\infty(\alpha^*(\delta))$ for some $\alpha^*(\delta) \in (0,1)$.*

PROOF: (Similar to Proposition 4. Can be omitted.) See Appendix.

5. CONCLUDING REMARKS

5.1. *Efficiency Wage and Three Types of Sanctions*

Our model describes a society where players meet a stranger to play a voluntarily separable prisoner's dilemma. We analyzed how continuous cooperation becomes an equilibrium behavior when deviation from cooperation induces appropriate social sanctions.

Sanctions consist of two parts. First, a player's defection invokes partner's severence decision, forcing him to start new partnership with a stranger. Second, payoff level he expects with this stranger is less than what he expects in continued partnership with the current partner. We call this payoff difference as *trust capital* with the ongoing partner.

In the main text, we have identified two ways by which trust is generated; positive trust-building periods and exploitation by strategies with longer trust-building periods.

There is an additional mechanism which creates trust if we allow matching probability to be less than one: Even if trust is established with new partner immediately, with a positive probability player fails to find a partner in the matching pool (i.e., player is "unemployed"). This is the logic which provides a work incentive in the effciency wage theory as the possibility of unemployment works as a disciplinary device (see, e.g., Shapiro and Stiglitz, 1984). For completeness of the paper we briefly discuss how our model can be extended to derive c_0-strategy as a symmetric NSD when there is a positive unemployment probability.

Suppose, in the matching pool, only with probability $1-u \in (0,1)$ one can find a new partner and with probability $u \in (0,1)$ he spends the next period without a partner and receives payoff of 0 (which may be larger or smaller than d). With this possibility of "unemployment", the average payoff that c_T-strategy player expects to receive in the matching pool (but before he finds a partner) is:

$$v^0(c_T; p_T) = (1-u)v(c_T; p_T),$$

where $v(c_T; p_T)$ is now interepreted as "the average payoff that c_T expects to receive when the new partnership is formed" (i.e., at the beginning of period 1 of the partnership).

By the same logic as in Section 3, the Best Reply Condition is $v^0(c_T; p_T) \leq v^{BR}$. Clearly, if (5) is satisfied, the Best Reply condition is also satisfied. Moreover, it can be satisfied even for c_0 for sufficiently large u, and cooperation without trust-building period becomes a self-sustaining state.[5]

As noted in Shapiro and Stiglitz (1984) and Okuno-Fujiwara (1989), unemployment works as a disciplinary device that deters moral hazard behavior. This observation suggests that the property of matching machanism is an important element in creating trust. In our setup, there are four reasons to be in the matching pool: new birth, death of the partner, separation due to the partner's deviation, and separation due to own deviation. In this paper we analyzed the case where no distinction can be made among these due to the lack of information. We plan to extend our research to investigate mechanisms with which players can distinguish at least some reasons why newly matched partner came into the matching pool.

5.2. *Alternating-Action Equilibrium*

If $2c < g + \ell$, then repeating (C, C) is not the most efficient outcome. It is most efficient to alternate (C, D) and (D, C). By a similar logic to the monomorphic

[5]Carmichael and MacLeod (1997) use essentially the same logic by gift-giving instead of unemployment.

equilibrium, the following two-strategy distribution constitutes a NSD for sufficiently long trust-building periods.

DEFINITION. For any $T = 1, 2, \ldots$, a_T-strategy is a simple strategy with the set of acceptable paths

$$\begin{aligned} Q_{aT} &= \{(\overbrace{(D,D),\cdots(D,D)}^{T\ \text{times}},(C,D),(D,C),\ldots), \\ &\quad (\overbrace{(D,D),\cdots(D,D)}^{T\ \text{times}},(C,C),(C,C),\ldots)\}. \end{aligned}$$

DEFINITION. For any $T = 1, 2, \ldots$, b_T-strategy is a simple strategy with the set of acceptable paths

$$\begin{aligned} Q_{bT} &= \{(\overbrace{(D,D),\cdots(D,D)}^{T\ \text{times}},(D,C),(C,D),\ldots), \\ &\quad (\overbrace{(D,D),\cdots(D,D)}^{T\ \text{times}},(D,D),(C,C),(C,C)\ldots),\}. \end{aligned}$$

If a_T met a_T, the play path is the same as c_T meeting c_T. If a_T met b_T, the play path after T periods of trust-building alternates action profiles (C, D) and (D, C). If b_T met b_T, the play path is the same as c_{T+1} meeting c_{T+1}.

Note that there is no voluntary separation on the play path even though there are multiple strategies in the society. Therefore the essential logic is the same as that of a monomorphic NSD. This type of equilibrium can be interpreted as a "single-norm" equilibrium with coordinated action profiles. The analysis will be useful for other types of component games of voluntarily separable games such as Hawk-Dove game, where the efficient outcome is a coordinated action profile.

5.3. *Cheap Talk*

Recall that c_1-strategy can invade the population of a $\tilde{d}$-strategy. We can interpret that c_1-strategy proposes to keep the partnership even after (D, D) and that this proposal acts as a "signal" or "cheap talk" that it is not $\tilde{d}$-strategy and intends to cooperate. This reminds us of papers like Robson (1990) and Matsui (1991) who showed that cheap talk can be used as a signal to play the Pareto efficient Nash equilibrium in coordination games. Because there are multiple NSD with different payoff outcomes in our model, cheap talk may work as a coordination device to achieve efficient equilibrium in evolutionary setting. We provide a rough sketch of what would happen if we allow cheap talk at the beginning of each match.

Assume that when two players are newly matched, they simultaneously choose and send a message $m \in M$ from a countable set M to the partner. M is common to all players. The messages do not alter the payoff and thus are cheap-talk. The message choice is private information, shared between the partners but not known by any other palyers.

Definition. A pure strategy s^{CT} of VSPD with cheap talk consists of (m, σ) such that:

1. $m \in M$ specifies the message the player sends to any new partner,
2. $\sigma : M \to \boldsymbol{S}$ specifies the VSPD strategy $\sigma(m')$ the player chooses to play for each message $m' \in M$ he receives from the partner.

Let $\boldsymbol{S}^{CT}$ be the set of all pure strategies of VSPD with cheap talk, which is the extension of $\boldsymbol{S}$ defined for the original VSPD without cheap talk.

We focus on two types of strategies; *babbling* strategy where message choice has no meaningful contents and *neologism* strategy where the message can be anything that is different from the ones used by the incumbents.

Definition. Given a strategy $s \in \boldsymbol{S}$ of VSPD, a strategy $s^B(s) = (m, \sigma^s) \in \boldsymbol{S}^{CT}$ of the cheap talk game is an *associated babbling strategy* of s if $\sigma^s(m') = s$ for all $m' \in M$.

Note that there is a class of associated babbling strategies of the same $s \in \boldsymbol{S}$ depending on the initial message m, but, if all players use associated babbling strategies of the same $s \in \boldsymbol{S}$, then the initial message does not matter. Thus we can focus on σ^s. Similarly, given a strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$, an associated babbling strategy distribution is denoted as $\sigma^p \in \mathcal{P}(\boldsymbol{S}^{CT})$. As is well-known, any babbling extension of a Nash equilibrium of the ordinary VSPD is always a Nash equilibrium of the cheap talk model because the initial message exchange does not matter.

LEMMA 13. *For any Nash Equilibrium $p \in \mathcal{P}(\boldsymbol{S})$ of VSPD, an associated babbling strategy distribution $\sigma^p \in \mathcal{P}(\boldsymbol{S}^{CT})$ is a Nash Equilibrium of the cheap talk model.*

PROOF: Obvious.

However, some babbling Nash equilibria are invaded by a neologism strategy in the cheap talk model. Suppose that the current population consists of babbling strategies of $s \in \boldsymbol{S}$. Against the strategy distribution, consider an entrant population who uses a strategy $s^N = (\zeta, \sigma^N) \in \boldsymbol{S}^{CT}$ such that

(a) it announces a neologism message $\zeta \in M$, which is not used by the current population,

(b) $\sigma^N(m') = s$ when $m' \neq \zeta$, and

(c) $\sigma^N(\zeta) = s' \neq s$.

With this neologism strategy, entrants play exactly the same way as incumbents (i.e., play s) when they are matched with incumbents, while they play differently (i.e., play according to s') against fellow entrants. They can identify incumbents who announce non-neologism messages from fellow entrants who announce neologism message at the initial message exchange. Therefore, for example, if the incumbents are playing associated babbling strategies of c_T-strategy, entrants can play c_{T-1} among themselves and earn higher average payoff than the incumbents. However, if the trust-building periods are shortened more and more, eventually the Best Reply Condition will be violated. Hence we require that entrants must be a best reply to the post-entry distribution, to avoid the non-existence of a stable distribution.[6]

[6]This idea is the same as Swinkels (1992).

Definition. A stationary strategy distribution p in the matching pool is a *Neutrally Stable Distribution under Equilibrium Entrants* (NSDEE) if, for any s', there exists $\bar{\epsilon} \in (0,1)$ such that, for any $\epsilon \in (0, \bar{\epsilon})$, s' is a best reply to $(1-\epsilon)p + \epsilon p_{s'}$ and s' cannot invade p for ϵ.

Let $s^* \in \boldsymbol{S}$ of VSPD be the most efficient NSS (i.e., the average payoff is the highest among NSS). Let $\sigma^{s^*} \in \mathcal{P}(S^{CT})$ be an associated babbling strategy distribution. Clearly, with cheap talk, no strategy can invade the most efficient NSS as an equlibrium entrant. Thus, we have the following result.

Proposition 6. *Among associated babbling strategy distributions of monomorphic NSDs, the most efficient $\sigma^{s^*} \in \mathcal{P}(S^{CT})$ is the unique NSDEE with cheap talk.*

Proof: Obvious.

5.4. *Drift and Limit of Solution Concept*

The concept of NSD is not sufficiently restrictive in our model, because any strategy distribution leaves many unreached nodes. Limitation of the concept of NSD is especially evident in view of the possibility of drift. As an example, consider the following thought experiment with or without cheap talk. Suppose G and δ are chosen so that c_1-strategy is the most efficient NSS. Since c_1 is NSDEE, once the entire society starts to use c_1 (or $\sigma^B(c_1)$ if cheap talk is allowed), no strategy can invade as an equilibrium entrant.

However, there are numerous strategies which produce exactly the same outcome (and hence the same average payoff) but differ in the behavior at unreached nodes. For example, consider the following strategy $\hat{c}_1 \in \boldsymbol{S}$:

$t = 1$: play D and choose k regardless of the outcome,

$t \geq 2$: play C and choose k regardless of the outcome.

This strategy produces exactly the same outcome as p_1 as long as the society consists only of c_1 and $\hat{c}_1$. Thus, strarting from p_1, strategy distribution may drift to any distribution $\gamma p_1 + (1-\gamma)\hat{p}_1$ with $\gamma \in [0,1]$, where $\hat{p}_1$ is the distribution consisting only of $\hat{c}_1$.

However, $\hat{c}_1$ being an extremely permissive strategy, strategies such as c_∞ can take advantage and materialize payoff stream of $(d, g, g, \dots)$ during the match with $\hat{c}_1$. Note that c_∞ can receive average payoff of only d in the strategy distribution p_1, which is strictly lower than that of c_1. However if drifts make γ sufficiently large, c_∞ starts to drive out c_1. Eventually, strategy distribution may become p_∞, the symmetric distribution consisting only of c_∞.

Such a story suggests that we might consider set-theoretic solution concepts, such as Equilibrium Evolutionary Stable Set of Swinkles (1992) or Socially Stable Strategies of Matsui (1992). In fact, drifts may lead from p_1 to $\hat{p}_1$, from $\hat{p}_1$ to p_∞, from p_∞ to $p_{\tilde{d}}$, and from $p_{\tilde{d}}$ back to p_1. However, there are many other closed paths which are connected by drifts (through equilibrium entrants). The cardinality of set of stragies being so large, we shall not try to identify these sets in this paper.

5.5. *Related Literature*

Several papers have previously analyzed the voluntarily separable games, though not as fully as this paper does. We discuss two main points of our paper in relation to the literature: the function of trust-building periods and the meaning of polymorphic equilibria.

First, the trust-building periods in our equilibria serve as a mechanism for sanction against defection because they make the initial value of a new partnership small. In the literature, the gift exchange of Carmichael and MacLeod (1997) and the gradual cooperation in Datta (1996) and Kranton (1996a) have the same function. By contrast, the gradual cooperation under incomplete information (Ghosh and Ray, 1996, and Kranton, 1996a) is to sort types out and thus has a different meaning.[7]

Our model is more primitive than these previous works: the game is of complete information, the component game is an ordinary prisoner's dilemma with two actions, and there is no gift exchange prior to the partnership. We show that it is still possible to construct a punishment mechanism. Furthermore, we consider

[7] The repeated games with quitting option (Watson, 2002, Blonski and Probst, 2001, and Furusawa and Kawakami, 2004) also display gradual cooperation to sort types.

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. We are also able to provide fuller characterizations of symmetric trust-building strategy NSD, such as indentifying the condition (in terms of death rate and payoffs of stage game) for the existence of NSD with a particular length of trust-building periods and so on.

Second, the existence and higher efficiency of polymorphic equilibria than monomorphic equilibria is a totally new result. The logic that early start of long-term cooperation is sustained because of possible exploitation in a new partnership is similar to the equilibrium of Rob and Yang (2005), written independently from our paper. 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's. Our model does not rely on heterogeous "type" and incomplete information. Instead, bad (longer trust-buidling) strategy emerges endogenously as a polymorphic NSD. We also show that there are equilibria with more than two (even infinitely many) heterogenous strategies.

APPENDIX: PROOFS

PROOF OF LEMMA 3:

(a) Let s' be a strategy that chooses e in some t after on-path history. If $t < T+1$, the average payoff of s' under p_T is d and is strictly less than $v(c_T; p_T) = (1-\delta^{2T})d+\delta^{2T}c$. If $t \geq T+1$, the average value is

$$\begin{aligned} L(s', c_T) &= \frac{1-\delta^{2t}}{1-\delta^2}, \\ V^I(s', c_T) &= (1+\delta^2+\cdots+\delta^{2(T-1)})d + (\delta^{2T}+\cdots+\delta^{2(t-1)})c, \\ v(s'; p_T) &= v^I(s', c_T) = \frac{1-\delta^2}{1-\delta^{2t}}\Big[\frac{1-\delta^{2T}}{1-\delta^2}d + \frac{\delta^{2T}(1-\delta^{2(t-T)})}{1-\delta^2}c\Big]. \end{aligned}$$

By computation,

$$\begin{aligned}
& \{v(c_T;p_T) - v(s';p_T)\}(1-\delta^{2t}) \\
= \;& (1-\delta^{2t})(1-\delta^{2T})d - (1-\delta^{2T})d + (1-\delta^{2t})\delta^{2T}c - \delta^{2T}(1-\delta^{2(t-T)})c \\
= \;& (1-\delta^{2T})\delta^{2t}(c-d) > 0.
\end{aligned}$$

(b) If one chooses C in $t < T+1$ along on-path history, then the average payoff is less than d since the partnership ends there and hence is less than $v(c_T;p_T) = (1-\delta^{2T})d + \delta^{2T}c$.

(c) Although the text contains a proof with one-step deviation argument, we provide an alternative proof using the average payoff itself to confirm that one-step deviation method is necessary and sufficient. Let s be any strategy that chooses D at some $t \geq T+1$ along on-path history.

$$\begin{aligned}
L(s, c_T) &= \frac{1-\delta^{2t}}{1-\delta^2}, \\
V^I(s, c_T) &= (1+\delta^2+\cdots+\delta^{2(T-1)})d + (\delta^{2T}+\cdots+\delta^{2(t-2)})c + \delta^{2(t-1)}g, \\
v(s;p_T) &= \frac{1-\delta^2}{1-\delta^{2t}}\Big[\frac{1-\delta^{2T}}{1-\delta^2}d + \frac{\delta^{2T}(1-\delta^{2(t-T-1)})}{1-\delta^2}c + \delta^{2(t-1)}g\Big].
\end{aligned}$$

By computation,

$$\begin{aligned}
& \{v(c_T;p_T) - v(s;p_T)\}(1-\delta^{2t}) \\
= \;& (1-\delta^{2t})(1-\delta^{2T})d + (1-\delta^{2t})\delta^{2T}c \\
& -(1-\delta^{2T})d - (\delta^{2T}-\delta^{2(t-1)})c - (1-\delta^2)\delta^{2(t-1)}g, \\
= \;& -\delta^{2t}(1-\delta^{2T})d + \delta^{2(t-1)}(1-\delta^{2T+2})c - (1-\delta^2)\delta^{2(t-1)}g, \\
= \;& \delta^{2(t-1)}\Big[-\delta^2(1-\delta^{2T})d + (1-\delta^2+\delta^2-\delta^{2T+2})c - (1-\delta^2)g\Big], \\
= \;& \delta^{2(t-1)}\Big[\delta^2(1-\delta^{2T})(c-d) - (1-\delta^2)(g-c)\Big].
\end{aligned}$$

Therefore $v(c_T;p_T) - v(s;p_T) \geq 0$ if and only if $\delta^2\frac{1-\delta^{2T}}{1-\delta^2}(c-d) \geq g-c$. Q.E.D.

PROOF OF LEMMA 4: Let $q := (1-\epsilon)p + \epsilon p_{s'}$. From (1), for any $s \in \text{supp}(p)$,

$$\begin{aligned}
v(s';q) &= (1-\epsilon)\frac{L(s';p)}{L(s';q)}v(s';p) + \epsilon\frac{L(s',s')}{L(s';q)}v^I(s',s'), \\
v(s;q) &= (1-\epsilon)\frac{L(s;p)}{L(s;q)}v(s;p) + \epsilon\frac{L(s,s')}{L(s;q)}v^I(s,s').
\end{aligned}$$

If s' invades p for some $\epsilon > 0$, then for any $s \in supp(p)$,

$$(1-\epsilon)\frac{L(s';p)}{L(s';q)}v(s';p) + \epsilon\frac{L(s',s')}{L(s';q)}v^I(s',s') \geq (1-\epsilon)\frac{L(s;p)}{L(s;q)}v(s;p) + \epsilon\frac{L(s,s')}{L(s;q)}v^I(s,s'),$$

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

$$(1-\epsilon)\frac{L(s';p)}{L(s';q)}v(s';p) + \epsilon\frac{L(s',s')}{L(s';q)}v^I(s',s') > (1-\epsilon)\frac{L(s;p)}{L(s;q)}v(s;p) + \epsilon\frac{L(s,s')}{L(s;q)}v^I(s,s').$$

By letting $\epsilon \to 0$, we obtain

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

for any $s \in supp(p)$. Since p is a Nash equilibrium, we have that $s' \in BR(p)$. Q.E.D.

PROOF OF LEMMA 5: Let us rearrange $v(c_T; p_T^{T+1}(\alpha))$ to highlight the effect of α.

$$\begin{aligned}
& v(c_T; p_T^{T+1}(\alpha)) \\
= & \frac{\alpha L(c_T, c_T) v^I(c_T, c_T) + (1-\alpha) L(c_T, c_{T+1}) v^I(c_T, c_{T+1})}{\alpha L(c_T, c_T) + (1-\alpha) L(c_T, c_{T+1})} \\
= & \frac{\alpha L(c_T, c_T) v^I(c_T, c_T)}{\alpha L(c_T, c_T) + (1-\alpha) L(c_T, c_{T+1})} + \frac{(1-\alpha) L(c_T, c_{T+1}) v^I(c_T, c_{T+1})}{\alpha L(c_T, c_T) + (1-\alpha) L(c_T, c_{T+1})} \\
= & \frac{\alpha L(c_T, c_T)}{\alpha L(c_T, c_T) + (1-\alpha) L(c_T, c_{T+1})} v^I(c_T, c_T) \\
& \quad + \Big[1 - \frac{\alpha L(c_T, c_T)}{\alpha L(c_T, c_T) + (1-\alpha) L(c_T, c_{T+1})}\Big] v^I(c_T, c_{T+1}) \\
= & v^I(c_T, c_{T+1}) + \frac{\alpha L(c_T, c_T)}{L(c_T, c_{T+1}) + \alpha\{L(c_T, c_T) - L(c_T, c_{T+1})\}} \{v^I(c_T, c_T) - v^I(c_T, c_{T+1})\}.
\end{aligned}$$

Let

$$\mu(c_T, p_T^{T+1}(\alpha)) := \frac{\alpha L(c_T, c_T)}{L(c_T, c_{T+1}) + \alpha\{L(c_T, c_T) - L(c_T, c_{T+1})\}}.$$

This is the only part that α is involved in $v(c_T; p_T^{T+1}(\alpha))$. We can simplfy as

$$v(c_T; p_T^{T+1}(\alpha)) = v^I(c_T, c_{T+1}) + \mu(c_T, p_T^{T+1}(\alpha)) \{v^I(c_T, c_T) - v^I(c_T, c_{T+1})\}. \tag{16}$$

By differentiation,

$$\frac{\partial \mu(c_T, p_T^{T+1}(\alpha))}{\partial \alpha} = \frac{L(c_T, c_T) L(c_T, c_{T+1})}{[L(c_T, c_{T+1}) + \alpha\{L(c_T, c_T) - L(c_T, c_{T+1})\}]^2} > 0,$$

and, since $L(c_T, c_T) - L(c_T, c_{T+1}) = \frac{1}{1-\delta^2} - \frac{1-\delta^{2(T+1)}}{1-\delta^2} > 0$, the derivative is decreasing in α. Note also that

$$\begin{aligned}
& v^I(c_T, c_T) - v^I(c_T, c_{T+1}) \\
= & (1-\delta^{2T}) d + \delta^{2T} c - \frac{(1-\delta^{2T}) d + \delta^{2T}(1-\delta^2)\ell}{1-\delta^{2(T+1)}} \\
= & \frac{(1-\delta^{2T})\{1-\delta^{2(T+1)} - 1\} d + \delta^{2T}\{(1-\delta^{2(T+1)}) c - (1-\delta^2)\ell\}}{1-\delta^{2(T+1)}} \\
= & \frac{\delta^{2T}\{(1-\delta^2)(c-\ell) + \delta^2(1-\delta^{2T})(c-d)\}}{1-\delta^{2(T+1)}} > 0.
\end{aligned}$$

Hence $v(c_T, p_T^{T+1}(\alpha))$ is strictly increasing and concave in α. Q.E.D

PROOF OF LEMMA 6:

$$\begin{aligned}
& v(c_{T+1}; p_T^{T+1}(\alpha)) \\
= & \frac{\alpha L(c_{T+1}, c_T) v^I(c_{T+1}, c_T) + (1-\alpha) L(c_{T+1}, c_{T+1}) v^I(c_{T+1}, c_{T+1})}{\alpha L(c_{T+1}, c_T) + (1-\alpha) L(c_{T+1}, c_{T+1})} \\
= & v^I(c_{T+1}, c_{T+1}) \\
& + \frac{\alpha L(c_{T+1}, c_T)}{L(c_{T+1}, c_{T+1}) + \alpha\{L(c_{T+1}, c_T) - L(c_{T+1}, c_{T+1})\}} \{v^I(c_{T+1}, c_T) - v^I(c_{T+1}, c_{T+1})\}.
\end{aligned}$$

Let

$$\mu(c_{T+1}, p_T^{T+1}(\alpha)) := \frac{\alpha L(c_{T+1}, c_T)}{L(c_{T+1}, c_{T+1}) + \alpha\{L(c_{T+1}, c_T) - L(c_{T+1}, c_{T+1})\}}.$$

Then

$$v(c_{T+1}; p_T^{T+1}(\alpha)) = v^I(c_{T+1}, c_{T+1}) + \mu(c_{T+1}, p_T^{T+1}(\alpha))\{v^I(c_{T+1}, c_T) - v^I(c_{T+1}, c_{T+1})\}. \tag{17}$$

Note that

$$\begin{aligned} & v^I(c_{T+1}, c_T) - v^I(c_{T+1}, c_{T+1}) \\ = \ & \{v^I(c_{T+1}, c_T) - v^I(c_T, c_T)\} + \{v^I(c_T, c_T) - v^I(c_{T+1}, c_{T+1})\} > 0, \end{aligned}$$

since $T \leq \underline{\tau}(\delta)$ (thus the first bracket is nonnegative) and c_T starts cooperation earlier than c_{T+1} (thus the second bracket is positive).

By differentiation,

$$\frac{\partial \mu(c_{T+1}, p_T^{T+1}(\alpha))}{\partial \alpha} = \frac{L(c_{T+1}, c_T)L(c_{T+1}, c_{T+1})}{[L(c_{T+1}, c_{T+1}) + \alpha\{L(c_{T+1}, c_T) - L(c_{T+1}, c_{T+1})\}]^2} > 0.$$

However, notice that $L(c_{T+1}, c_T) - L(c_{T+1}, c_{T+1}) = \frac{1-\delta^{2(T+1)}}{1-\delta^2} - \frac{1}{1-\delta^2} < 0$ so that the derivative is increasing in α. Therefore $v(c_{T+1}; p_T^{T+1}(\alpha))$ is strictly increasing but convex in α. Q.E.D

PROOF OF LEMMA 7: Let $\mu_T(\alpha) = \frac{\alpha L(c_T, c_T)}{L(c_T; p_T^{T+1}(\alpha))}$ and $\mu_{T+1}(\alpha) = \frac{\alpha L(c_{T+1}, c_T)}{L(c_{T+1}; p_T^{T+1}(\alpha))}$. Then

$$\begin{aligned} v(c_T; p_T^{T+1}(\alpha)) &= \mu_T(\alpha) v^I(c_T, c_T) + \{1 - \mu_T(\alpha)\} v^I(c_T, c_{T+1}), \\ v(c_{T+1}; p_T^{T+1}(\alpha)) &= \mu_{T+1}(\alpha) v^I(c_{T+1}, c_T) + \{1 - \mu_{T+1}(\alpha)\} v^I(c_{T+1}, c_{T+1}). \end{aligned}$$

By differentiation,

$$\begin{aligned} \frac{\partial v(c_T; p_T^{T+1}(\alpha))}{\partial \alpha} &= \mu'_T(\alpha)\{v^I(c_T, c_T) - v^I(c_T, c_{T+1})\}, \\ \frac{\partial v(c_{T+1}; p_T^{T+1}(\alpha))}{\partial \alpha} &= \mu'_{T+1}(\alpha)\{v^I(c_{T+1}, c_T) - v^I(c_{T+1}, c_{T+1})\}. \end{aligned}$$

As $\alpha \to 1$,

$$\begin{aligned} \mu'_T(\alpha) &= \frac{L(c_T, c_T)L(c_T, c_{T+1})}{[\alpha L(c_T, c_T) + (1-\alpha)L(c_T, c_{T+1}))]^2} \to \frac{L(c_T, c_{T+1}))}{L(c_T, c_T)} = 1 - \delta^{2(T+1)}, \\ \mu'_{T+1}(\alpha) &= \frac{L(c_{T+1}, c_T)L(c_T, c_{T+1})}{[\alpha L(c_{T+1}, c_T) + (1-\alpha)L(c_{T+1}, c_{T+1}))]^2} \\ &\to \frac{L(c_{T+1}, c_{T+1})}{L(c_{T+1}, c_T)} = \frac{L(c_T, c_T)}{L(c_T, c_{T+1})} = \frac{1}{1 - \delta^{2(T+1)}}. \end{aligned}$$

At $\delta = \underline{\delta}(T)$,

$$v(c_T; p_T^{T+1}(1)) = v^I(c_T, c_T) = v(c_{T+1}; p_T^{T+1}(1)) = v^I(c_{T+1}, c_T).$$

Therefore, at $\delta = \underline{\delta}(T)$,

$$
\begin{aligned}
& \frac{\partial v(c_T; p_T^{T+1}(\alpha))}{\partial \alpha}\Big|_{\alpha=1} - \frac{\partial v(c_{T+1}; p_T^{T+1}(\alpha))}{\partial \alpha}\Big|_{\alpha=1} \\
= \; & \frac{L(c_T, c_{T+1})}{L(c_T, c_T)}\{v^I(c_T, c_T) - v^I(c_T, c_{T+1})\} \\
& -\frac{L(c_T, c_T)}{L(c_T, c_{T+1})}\{v^I(c_{T+1}, c_T) - v^I(c_{T+1}, c_{T+1})\}, \\
= \; & (1-\delta^{2(T+1)})\frac{\delta^{2T}(1-\delta^2)(g-\ell)}{1-\delta^{2(T+1)}} - \frac{1}{1-\delta^{2(T+1)}}\delta^{2T}(1-\delta^2)(c-d) \\
= \; & \delta^{2T}(1-\delta^2)\Big\{(g-\ell) - \frac{c-d}{1-\delta^{2(T+1)}}\Big\}.
\end{aligned}
$$

Q.E.D.

PROOF OF LEMMA 9: We consider the continuation average values of the incumbent strategies (c_T and c_{T+1}) and one-step deviant strategies. It suffices to check two kinds of one-step deviations during the cooperation periods of either c_T or c_{T+1}. (Note that unlike the monomorphic case, you cannot guarantee to get g at $T+1$ by playing D. Therefore we cannot use the computation in the monomorphic case directly.)

1. Imitate c_T in the first T periods. For $t \geq T+1$, play D and keep the partnership regardless of the outcome, until your partner ends the partnership.

   Using the continuation values, this type of deviation is not better than c_{T+1}-strategy if and only if

$$
\begin{aligned}
& \alpha\{g + \delta V(c_T; p_T^{T+1}(\alpha))\} \\
& \quad +(1-\alpha)\{d + \delta(1-\delta)V(c_T; p_T^{T+1}(\alpha)) + \delta^2 g + \delta^3 V(c_T; p_T^{T+1}(\alpha))\} \\
& \leq \alpha\{g + \delta V(c_T; p_T^{T+1}(\alpha))\} + (1-\alpha)\{d + \frac{\delta(1-\delta)}{1-\delta^2}V(c_T; p_T^{T+1}(\alpha)) + \frac{\delta^2 c}{1-\delta^2}\} \\
\iff \quad & v(c_T; p_T^{T+1}(\alpha)) = (1-\delta)V(c_T; p_T^{T+1}(\alpha)) \leq \frac{c-(1-\delta^2)g}{\delta^2} = v^{BR}
\end{aligned}
$$

2. Imitate c_T in the first T periods. For $t = T+1$, play C and keep the partnership regardless of the outcome. At $t = T+2$, play D and keep the partnership regardless of the outcome. Denote a strategy in this class by s_T.

   This type of deviation is not better than c_T-strategy if and only if

$$
\begin{aligned}
& \alpha\{c + \delta(1-\delta)V(c_T; p_T^{T+1}(\alpha)) + \delta^2 g + \delta^3 V(c_T; p_T^{T+1}(\alpha))\} \\
& \quad +(1-\alpha)\{\ell + \delta V(c_T; p_T^{T+1}(\alpha))\} \\
& \leq \alpha\{\frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2}V(c_T; p_T^{T+1}(\alpha))\} + (1-\alpha)\{\ell + \delta V(c_T; p_T^{T+1}(\alpha))\} \\
\iff \quad & v(c_T; p_T^{T+1}(\alpha)) = (1-\delta)V(c_T; p_T^{T+1}(\alpha)) \leq \frac{c-(1-\delta^2)g}{\delta^2} = v^{BR}.
\end{aligned}
$$

Q.E.D.

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

$$\Gamma(c_T, c_{T'}) = d - v^{BR} + \delta^2 \Gamma(c_{T-1}, c_{T'-1}). \tag{18}$$

Proof of Lemma 14: 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 15. *For any $T \in \mathbb{N}$ and for any $v \in \mathbb{R}$:*

$$L(c_{T+1}, c_T)\{v^I(c_{T+1}, c_T) - v\} \gtreqless L(c_T, c_T)\{v^I(c_T, c_T) - v\} \iff v \gtreqless v^{BR}. \tag{19}$$

Proof of Lemma 15: 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 14,

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

and the last inequalities hold by the induction assumption. $\square$

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. □

Because of the concavity of $v(c_T; p_T^{T+1}(\alpha))$ and convexity of $v(c_{T+1}; p_T^{T+1}(\alpha))$, thanks to Corollary 2 and continuity of average values with respect to T, the next lemma is immediate. (See Figure 3.)

LEMMA 16. *For any* $\delta > \delta^*$*, there exists* $0 \leq \tau_2(\delta) < \underline{\tau}(\delta)$ *such that, if* $\tau_2(\delta) < T < \underline{\tau}(\delta)$,

(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 $\tau_2(\delta) < T < \underline{\tau}(\delta)$, 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

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

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

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

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

Hence by computation,

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

Therefore the Best Reply Condition is satisfied. Q.E.D.

PROOF OF LEMMA 10: By the definition of $\alpha^*_{T+1}(v^{BR})$,

$$\begin{aligned} & v(c_{T+1}; p^{T+1}_T(\alpha^*_{T+1}(v^{BR})) = v^{BR}, \\ \Longleftrightarrow \; & \alpha^*_{T+1}(v^{BR})L(c_{T+1}, c_T)v^I(c_{T+1}, c_T) + \{1 - \alpha^*_{T+1}(v^{BR})\}L(c_{T+1}, c_{T+1})v^I(c_{T+1}, c_{T+1}) \\ & = \alpha^*_{T+1}(v^{BR})L(c_{T+1}, c_T)v^{BR} + \{1 - \alpha^*_{T+1}(v^{BR})\}L(c_{T+1}, c_{T+1})v^{BR}, \\ \Longleftrightarrow \; & \alpha^*_{T+1}(v^{BR})\Gamma(c_{T+1}, c_T) + \{1 - \alpha^*_{T+1}(v^{BR})\}\Gamma(c_{T+1}, c_{T+1}) = 0. \end{aligned}$$

By Corollary 1

$$\begin{aligned} \Longleftrightarrow \; & \alpha^*_{T+1}(v^{BR})\Gamma(c_{T+1}, c_T) + \{1 - \alpha^*_{T+1}(v^{BR})\}\Gamma(c_{T+1}, c_{T+1}) = 0, \\ \Longleftrightarrow \; & v(c_{T+2}; p^{T+2}_T(\alpha^*_{T+1}(v^{BR}), 1)) = v^{BR}, \\ \Longleftrightarrow \; & v(c_{T+1}; p^{T+2}_T(\alpha^*_{T+1}(v^{BR}), 1)) = v^{BR}. \end{aligned}$$

Next, we prove that for any $\alpha < \alpha^*_{T+1}(v^{BR})$,

$$v(c_{T+1}; p^{T+2}_T(\alpha, \beta^*(\alpha))) < v(c_{T+1}; p^{T+1}_T(\alpha)).$$

For any (α, v), define

$$\begin{aligned} \Phi^*_{T+1}(\alpha, v) &:= L(c_{T+1}; p^{T+1}_T(\alpha))\{v(c_{T+1}; p^{T+1}_T(\alpha)) - v\} \\ &= \alpha\Gamma(c_{T+1}, c_T) + (1-\alpha)\Gamma(c_{T+1}, c_{T+1}) \\ &\quad \{\alpha L(c_{T+1}, c_T) + (1-\alpha)L(c_{T+1}, c_{T+1})\}(v^{BR} - v), \\ \Phi^{**}_{T+1}(\alpha, v) &:= L(c_{T+1}; p^{T+2}_T(\alpha, \beta^*(\alpha)))\{v(c_{T+1}; p^{T+2}_T(\alpha, \beta^*(\alpha))) - v\} \\ &= \alpha\Gamma(c_{T+1}, c_T) + (1-\alpha)\{\beta^*(\alpha)\Gamma(c_{T+1}, c_{T+1}) + (1-\beta^*(\alpha))\Gamma(c_{T+1}, c_{T+2}) \\ &\quad +[\alpha L(c_{T+1}, c_T) + (1-\alpha)\{\beta^*(\alpha)L(c_{T+1}, c_{T+1}) \\ &\qquad +(1-\beta^*(\alpha))L(c_{T+1}, c_{T+2})\}](v^{BR} - v). \end{aligned}$$

Then

$$\begin{aligned}&\Phi^*_{T+1}(\alpha, v) - \Phi^{**}_{T+1}(\alpha, v)\\ = \;&(1-\alpha)(1-\beta^*(\alpha))[\{\Gamma(c_{T+1}, c_{T+1}) - \Gamma(c_{T+1}, c_{T+2})\}\\ &+\{L(c_{T+1}, c_{T+1}) - L(c_{T+1}, c_{T+2})\}(v^{BR} - v)].\end{aligned}$$

By computation

$$\{\Gamma(c_{T+1}, c_{T+1}) - \Gamma(c_{T+1}, c_{T+2})\}(1-\delta^2) = \delta^{2(T+1)}\{\delta^2(c - v^{BR}) + (1-\delta^2)(c-\ell)\} > 0.$$

Hence, $\Phi^*_{T+1}(\alpha, v) > \Phi^{**}_{T+1}(\alpha, v)$ if $v \leq v^{BR}$. Now,

$$\begin{aligned}&L(c_{T+1}; p_T^{T+2}(\alpha, \beta^*(\alpha))[\{v(c_{T+1}; p_T^{T+1}(\alpha)) - v\} - \{v(c_{T+1}; p_T^{T+2}(\alpha)) - v\}]\\ = \;&\Phi^*_{T+1}(\alpha, v) - \Phi^{**}_{T+1}(\alpha, v)\\ &-(1-\alpha)(1-\beta^*(\alpha))\{L(c_{T+1}, c_{T+1}) - L(c_{T+1}, c_{T+2})\}\{v(c_{T+1}; p_T^{T+1}(\alpha)) - v\}.\end{aligned}$$

Let $v = v(c_{T+1}; p_T^{T+1}(\alpha))$, then the above implies that

$$v(c_{T+1}; p_T^{T+2}(\alpha, \beta^*(\alpha))) < v(c_{T+1}; p_T^{T+1}(\alpha))$$

if $v(c_{T+1}; p_T^{T+1}(\alpha)) < v^{BR}$. Q.E.D.

PROOF OF LEMMA 11: Consider c_t-strategy for an arbitrary $t \in \{T, T+1, T+2, \ldots\}$ and the beginning of period $t+1$ in a match, when c_t-strategy is about to start cooperation. Let α_t be the conditional probability that the partner is the same strategy. The conditional probability is $1-\alpha_t$ that the partner has a longer trust-building period. The (non-averaged) continuation payoff of c_t-strategy at the beginning of $t+1$ is

$$V(c_t; p, t+1) = \alpha_t\{\frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2}V(c_t; p)\} + (1-\alpha_t)\{\ell + \delta V(c_t; p)\}. \tag{20}$$

On the other hand, the continuation payoff of c_{t+1}-strategy is

$$\begin{aligned}V(c_{t+1}; p, t+1) \;=\; &\alpha_t\{g + \delta V(c_{t+1}; p)\}\\ &+(1-\alpha_t)\{d + \delta(1-\delta)V(c_{t+1}; p) + \delta^2 V(c_{t+1}; p, t+2)\}.\end{aligned} \tag{21}$$

Notice that the payoff structure for c_{t+1}-strategy at the beginning of period $t+2$ when it just finished the trust building is the same as that of c_t-strategy at $t+1$, i.e.,

$$V(c_{t+1}; p, t+2) = V(c_t; p, t+1).$$

Therefore (21) becomes

$$\begin{aligned}V(c_{t+1}; p, t+1) \;=\; &\alpha_t\{g + \delta V(c_{t+1}; p)\}\\ &+(1-\alpha_t)\{d + \delta(1-\delta)V(c_{t+1}; p) + \delta^2 V(c_t; p, t+1)\}\\ \iff V(c_{t+1}; p, t+1) \;=\; &\frac{1}{1-(1-\alpha_t)\delta^2}\big[\alpha_t\{g + \delta V(c_{t+1}; p)\}\\ &+(1-\alpha_t)\{d + \delta(1-\delta)V(c_{t+1}; p)\}\big].\end{aligned} \tag{22}$$

From the assumption that the average payoffs of c_t and c_{t+1} are the same,

$$V(c_t; p) = V(c_{t+1}; p). \tag{23}$$

Then, since the payoff until t is the same for both c_t and c_{t+1}, we also have

$$V(c_t; p, t+1) = V(c_{t+1}; p, t+1). \tag{24}$$

(24) implies that the RHS of (20) and (22) must be the same. Using (23) and letting $V^*(p) = V(c_t; p) = V(c_{t+1}; p)$, α_t must satisfy

$$\begin{aligned} & \alpha_t\{\frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2}V^*(p)\} + (1-\alpha_t)\{\ell + \delta V^*(p)\} \\ = \quad & \frac{\alpha_t\{g + \delta V^*(p)\} + (1-\alpha_t)\{d + \delta(1-\delta)V^*(p)\}}{1-(1-\alpha_t)\delta^2}. \end{aligned}$$

Since this equation does not depend on t, we have established that $\alpha_t = \alpha$ for all $t = T, T+1, \ldots$, i.e., the fraction of $c_{T+\tau}$-strategy is of the form $\alpha(1-\alpha)^\tau$. Q.E.D.

PROOF OF LEMMA 12: For any strategy pair (s, s') and any stationary distribution p in the matching pool, let $V(s, s'; p)$ be the (non-averaged) payoff of strategy s when it is newly matched with s', i.e.,

$$V(s, s'; p) = V^I(s, s') + \{1-(1-\delta)L(s, s')\}V(s; p),$$

where the first term of the RHS is the in-match payoff and the second term is the expected payoff when s-strategy looses the partner either by the death or because they reached the ending date $T(s, s')$. (See Section 2.3.) Then the long-run payoff of c_T-strategy is decomposed as

$$\begin{aligned} V(c_T; p_T^\infty(\alpha)) \quad = \quad & \alpha V(c_T, c_T; p_T^\infty(\alpha)) \\ & +(1-\alpha)V(c_T, c_{T+1}; p_T^\infty(\alpha)). \end{aligned} \tag{25}$$

The long-run payoff of c_{T+1}-strategy is decomposed as

$$\begin{aligned} V(c_{T+1}; p_T^\infty(\alpha)) \quad = \quad & \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) \\ & +(1-\alpha)[\alpha\{d + \delta^2 V(c_T, c_T; p_T^\infty(\alpha)) + \delta(1-\delta)V(c_{T+1}; p_T^\infty(\alpha))\} \\ & \qquad (1-\alpha)\{d + \delta^2 V(c_T, c_{T+1}; p_T^\infty(\alpha)) + \delta(1-\delta)V(c_{T+1}; p_T^\infty(\alpha))\}] \\ = \quad & \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) \\ & +(1-\alpha)[d + \delta^2 V(c_T; p_T^\infty(\alpha)) + \delta(1-\delta)V(c_{T+1}; p_T^\infty(\alpha))], \end{aligned} \tag{26}$$

where the last equality uses (25). The intuition is easily understood from Table II(b). The equality (26) is equivalent to

$$\begin{aligned} [1-(1-\alpha)\delta(1-\delta)]V(c_{T+1}; p_T^\infty(\alpha)) \quad = \quad & \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) + (1-\alpha)d \\ & +(1-\alpha)\delta^2 V(c_T; p_T^\infty(\alpha)). \end{aligned} \tag{27}$$

Similarly from Table II(b) and II(c),

$$\begin{aligned} V(c_{T+2}; p_T^\infty(\alpha)) \quad = \quad & \alpha V(c_{T+2}, c_T; p_T^\infty(\alpha)) \\ & (1-\alpha)[d + \delta^2 V(c_{T+1}; p_T^\infty(\alpha)) + \delta(1-\delta)V(c_{T+2}; p_T^\infty(\alpha))]. \end{aligned}$$

TABLE II

(a): Payoff sequence of c_T-strategy under $p_T^\infty(\alpha)$ within a match

prob.	partner \ time	1	2	$\cdots$	T	$T+1$	$T+2$	$T+3$	$T+4$
α	c_T	d	d	$\cdots$	d	c	c	c	$\cdots$
$(1-\alpha)$	c_{T+1} and up	d	d	$\cdots$	d	ℓ			

(b): Payoff sequence of c_{T+1}-strategy under $p_T^\infty(\alpha)$ within a match

prob.	partner \ time	1	2	$\cdots$	T	$T+1$	$T+2$	$T+3$	$T+4$
α	c_T	d	d	$\cdots$	d	g			
$(1-\alpha)\alpha$	c_{T+1}	d	$\boldsymbol{d}$	$\cdots$	$\boldsymbol{d}$	$\boldsymbol{d}$	$\boldsymbol{c}$	$\boldsymbol{c}$	$\cdots$
$(1-\alpha)^2$	c_{T+2} and up	d	$\boldsymbol{d}$	$\cdots$	$\boldsymbol{d}$	$\boldsymbol{d}$	$\boldsymbol{\ell}$		

(c) : Payoff sequence of c_{T+2}-strategy under $p_T^\infty(\alpha)$ within a match

prob.	partner \ time	1	2	$\cdots$	T	$T+1$	$T+2$	$T+3$	$T+4$
α	c_T	d	d	$\cdots$	d	g			
$(1-\alpha)\alpha$	c_{T+1}	d	$\boldsymbol{d}$	$\cdots$	$\boldsymbol{d}$	$\boldsymbol{d}$	$\boldsymbol{g}$		
$(1-\alpha)^2\alpha$	c_{T+2}	d	$\boldsymbol{d}$	$\cdots$	$\boldsymbol{d}$	$\boldsymbol{d}$	$\boldsymbol{d}$	$\boldsymbol{c}$	$\cdots$
$(1-\alpha)^3$	c_{T+3} and up	d	$\boldsymbol{d}$	$\cdots$	$\boldsymbol{d}$	$\boldsymbol{d}$	$\boldsymbol{d}$	$\boldsymbol{\ell}$	

Note that c_{T+1} and c_{T+2} earn the same payoff against c_T and thus $V(c_{T+2}, c_T; p_T^\infty(\alpha)) = V(c_{T+1}, c_T; p_T^\infty(\alpha))$. Therefore the long-run payoff of c_{T+2}-strategy solves

$$\begin{aligned} V(c_{T+2}; p_T^\infty(\alpha)) &= \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) \\ &\quad +(1-\alpha)[[d + \delta^2 V(c_{T+1}; p_T^\infty(\alpha)) + \delta(1-\delta)V(c_{T+2}; p_T^\infty(\alpha))]. \end{aligned}$$

This is equivalent to

$$\begin{aligned} [1-(1-\alpha)\delta(1-\delta)]V(c_{T+2}; p_T^\infty(\alpha)) &= \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) + (1-\alpha)d \\ &\quad +(1-\alpha)\delta^2 V(c_{T+1}; p_T^\infty(\alpha)). \end{aligned} \tag{28}$$

If $V(c_T; p_T^\infty(\alpha)) = V(c_{T+1}; p_T^\infty(\alpha))$, then the last term of the right hand sides of (27) and (28) are the same and therefore

$$V(c_{T+1}; p_T^\infty(\alpha)) = V(c_{T+2}; p_T^\infty(\alpha)).$$

We can continue this argument for any $t > T$. Q.E.D.

PROOF OF PROPOSITION 5: We use the same logic as the proof of Proposition 4. Let

$$\begin{aligned} \Phi_{T+1}^\infty(\alpha, v) &:= L(c_{T+1}; p_T^\infty(\alpha))\{v(c_{T+1}; p_T^\infty(\alpha)) - v\} \\ &= \alpha\Gamma(c_{T+1}, c_T) + (1-\alpha)\alpha\Gamma(c_{T+1}, c_{T+1}) + (1-\alpha)^2\Gamma(c_{T+1}, c_{T+2}) \\ &\quad +\{\alpha L(c_{T+1}, c_T) + (1-\alpha)\alpha L(c_{T+1}, c_{T+1}) + (1-\alpha)^2 L(c_{T+1}, c_{T+2})\}(v^{BR} - v). \end{aligned}$$

Therefore

$$\begin{aligned} & \Phi^*_{T+1}(\alpha, v) - \Phi^\infty_{T+1}(\alpha, v) \\ = \ & (1-\alpha)^2[\Gamma(c_{T+1}, c_{T+1}) - \Gamma(c_{T+1}, c_{T+2}) \\ & +\{L(c_{T+1}, c_{T+1}) - L(c_{T+1}, c_{T+2})\}(v^{BR} - v) > 0, \end{aligned}$$

if $v \leq v^{BR}$. Now,

$$\begin{aligned} & L(c_{T+1}; p_T^{T+2}(\alpha, \beta^*(\alpha))[\{v(c_{T+1}; p_T^{T+1}(\alpha)) - v\} - \{v(c_{T+1}; p_T^\infty(\alpha)) - v\}] \\ = \ & \Phi^*_{T+1}(\alpha, v) - \Phi^\infty_{T+1}(\alpha, v) \\ & -(1-\alpha)^2\{L(c_{T+1}, c_{T+1}) - L(c_{T+1}, c_{T+2})\}\{v(c_{T+1}; p_T^{T+1}(\alpha)) - v\}. \end{aligned}$$

Let $v = v(c_{T+1}; p_T^{T+1}(\alpha))$, then the above implies that

$$v(c_{T+1}; p_T^\infty(\alpha, \beta^*(\alpha))) < v(c_{T+1}; p_T^{T+1}(\alpha))$$

if $v(c_{T+1}; p_T^{T+1}(\alpha)) < v^{BR}$. Hence if $v(c_{T+1}; p_T^{T+1}(\alpha))$ intersects with $v(c_T; p_T^{T+1}(\alpha))$ below v^{BR}, then so does $v(c_{T+1}; p_T^\infty(\alpha))$. Q.E.D.

REFERENCES

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

BLONSKI, M., and D. PROBST (2001): "The Emergence of Trust," mimeo., University of Mannheim.

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)

FURUSAWA,T., and T. KAWAKAMI (2004): "Gradual Cooperation in the Existence of Outside Options," mimeo., Hitotsubashi University.

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.

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

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

SWINKELS, J. (1992): "Evolutionary Stability with Equilibrium Entrants," *Journal of Economic Theory*, 57, 306-332.

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