Voluntarily Repeated 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

Still preliminary. Comments welcome.
This draft: December 11, 2005

Abstract: We consider a large society of homogeneous players, in which players are randomly matched to play prisoner's dilemma as well as to choose whether to play the game again with the same partner. Moreover, there is no information flow across matches because of private actions and random deaths. We extend the notion of Neutrally Stable Distribution (NSD) into our extensive-form model and characterize *simple-strategy* NSDs which end the partnership as soon as unacceptable action is observed. Even though personalized punishment is impossible, we can achieve eventual cooperation by initial *trust-building* periods and/or occasional exploitation by players with longer trust-building periods. Asymmetric strategy NSDs with voluntary break-ups on the equilibrium path has strategies with shorter trust-building periods than symmetric NSDs and becomes more efficient. When cheap-talk is introduced the most efficient NSD is the unique NSD that cannot be invaded by equilibrium entrants. (142 words)

Key words: voluntarily repeated prisoner's dilemma, evolution, efficiency, neutral stability, cheap-talk.
JEL classification number: C 73

*We thank useful comments from Nobue Suzuki, Akihiko Matsui, Chiaki Hara, and seminar participants at KIER, Kyoto University, Keio University, and University of Tokyo. All errors are ours. This work was supported by Grant-in-Aid for Scientific Research on Priority Areas, No.12124204.

1. INTRODUCTION

We consider a voluntarily repeated game in a large society of homogeneous players. Players are randomly matched to play a two-person prisoner's dilemma, 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. In a partnership, there is a merit of mutual cooperation but there is a gain by free-riding on the partner's cooperation as well. 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.

We consider boundedly rational players who are endowed with a pure strategy to play the voluntarily repeated game and develop a general framework of evolutionary stability that can be applied to general component games. A voluntarily repeated game is an extensive form game, and thus there are many strategies that only differ in the off-path. Hence invasion concept needs to be carefully defined. We extend Neutrally Stable Distribution (NSD) concept to our extensive form model, under which no other strategy earns strictly higher payoff than the incumbents do.

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

Previous literature on voluntarily repeated games focused on symmetric strategy distribution in which all (rational) players play the same strategy and showed that a gradual-cooperation strategy sustains eventual cooperation (Datta, 1996, Kranton 1996a, and Ghosh and Ray, 1997) for many-action prisoner's dilemma as the component game. By contrast, we consider the ordinary two-action prisoner's dilemma (and thus gradual increase in the cooperation level is not feasible) and both symmetric strategy distributions and asymmetric strategy distributions in which multiple strategies co-exist in the population. Among asymmetric strategy distributions, some can be considered as similar to a symmetric distribution in the sense that there is no voluntary break-up on the equilibrium path. (These are called *single-norm* strategy distributions.) Others are fundamentally asymmetric with voluntary break-ups, called *multi-norm* strategy distributions. Asymmetric strategy distributions among homogeneous players has not been considered in the literature.

To make the analysis as thorough as possible, we focus on *simple strategy* NSDs, consisting of simple strategies which ends the partnership as soon as the observed action path is not *acceptable*. Since a deviator can unilaterally end a partnership, no other punishment can achieve a lower continuation payoff than ending the partnership. Moreover, simple strategies have natural interpretation of social behavior of boundedly rational players, following a *norm* and escaping from norm-violators.

Single-norm NSDs require sufficiently long periods of *trust-building* such that partners play (D, D) but keep the partnership. The logic is the same as the gradual cooperation literature that deviators are punished by a low payoff of new partnerships to start trust-building again. We identify a relationship between the death rate (discount factor) and the sufficient length of trust-building periods.

By contrast, *multi-norm* NSDs can start the cooperation periods earlier than single-norm NSDs, thanks to possible exploitation by another strategy, which also serves as a punishment. Hence under the same payoff parameters and death rate, multi-norm NSD is more efficient than the most efficient single-norm NSD. The idea that diverse strategies make it valuable to keep a partnership with the same-strategy partner is similar to Ghosh and Ray (1997) and Rob and Yang (2005) of incomplete information model. We show that diverse strategies arise in a complete information, homogeneous population game.

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 NSD is the unique NSD that cannot be invaded by equilibrium entrants (Swinkels, 1992).

This paper is organized as follows. In Section 2, we introduce the formal model and stability concepts. In Section 3, we give necessary and sufficient conditions for single-norm NSD. In Section 4, we analyze multi-norm NSD. 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.

When a player is newly born, he enters into the *matching pool* where players are randomly paired to play a *Voluntarily Repeated Prisoner's Dilemma (VRPD)* as follows.

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 and be randomly paired to play another VRPD anew. 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 VRPD 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 with each play action profile as: $u(C, C) = c$, $u(C, D) = \ell$, $u(D, C) = g$, $u(D, D) = d$ with the ordering $g > c > d > \ell$. (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 VRPD) 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.[1] Let

$$H_t := \{C, D\}^{2(t-1)}$$

be *the set of partnership histories at the beginning of* $t \geq 2$ and let $H_1 := \{\emptyset\}$.

TABLE I

PAYOFF OF PD

P1 \ P2	C	D
C	c, c	ℓ, g
D	g, ℓ	d, d

[1]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.

Definition. A *pure strategy* s of VRPD specifies $(x_t, y_t)_{t=1}^{\infty}$ where:
$x_t : H_t \rightarrow \{C, D\}$ specifies an action choice $x_t(h_t) \in \{C, D\}$ given the partnership history $h_t \in H_t$, and
$y_t : H_t \times \{C, D\}^2 \rightarrow \{k, e\}$ specifies whether or not the player wants to keep or end the partnership, depending upon the partnership history $h_t \in H_t$ at the beginning of t and the current period action profile.

The (infinite) set of pure strategies of VRPD is denoted as $\boldsymbol{S}$ and the set of all strategy distributions in the population is denoted as $\mathcal{P}(\boldsymbol{S})$. For simplicity we assume that each player uses a pure strategy.

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

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 lifetime 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

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

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,

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

Hence the average payoff can be decomposed[2] 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 *monomorphic strategy distribution*[3]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 $s' \in \boldsymbol{S}$,

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

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 $s \in \text{supp}(p)$,

$$s \in BR(p).$$

For any pure strategy $s \in S$, let p_s be the ("monomorphic") strategy distribution consisting only of s.

LEMMA 1. *For any pure strategy $s \in S$ that starts with C in $t = 1$, p_s is not a Nash equilibrium.*

---

[2]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.

[3]If the partnership outcome is monomorphic, i.e., the same for all matches in the society, then we call a strategy distribution as monomorphic *outcome* distribution. This distinction becomes important in Section 5 where we consider equivalent strategies.

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

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

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

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

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 equilibrium.

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

2.5. *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 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 monomorphic 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 monomorphic 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.

An underlying assumption of this definition is that mutation takes place rarely so that only single mutation occurs within the time span in which stationary strategy distribution is formed. To compute the average payoff we need the expected length of the life-time of a mutant strategy. Our definition requires 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.

The next proposition shows that the myopic $\tilde{d}$-strategy can be invaded by a *trust-building* strategy which initially plays D but keeps the partnership to distinguish itself from the myopic strategy. If it meets the same strategy, they can play C after the first period, and if it meets the myopic strategy, it does not earn worse payoff than the myopic strategy.

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

PROOF: Consider the following c_1-strategy.

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

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

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

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

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

3. SINGLE-NORM, TRUST-BUILDING STRATEGY DISTRIBUTIONS

The c_1-strategy that can invade the myopic $\tilde{d}$-strategy distribution has the property that it keeps the partnership if and only if the sequence of action profiles $\{(D, D), (C, C), (C, C), \ldots\}$ is followed in the partnership. One can interpret the strategy as having the path $\{(D, D), (C, C), (C, C), \ldots\}$ as its "norm" and punishing as soon as the norm is violated.

In the following, we restrict our attention to such strategies, to make our analysis as complete as possible. S contains infinitely many strategies, some of which only differ in off-path actions, and it is endless to try to characterize all possible equilibria.

Specifically, we focus on the following *simple strategies*, which carry out the maximal individually rational punishment to end a partnership if unexpected behavior is observed. The idea of simple strategies is similar to that of Abreu's (1988), except that we define strategies instead of strategy profiles to fit for our evolutionary model. There is a set of *acceptable paths* of action profiles for a player/strategy, and if there is a deviation from the acceptable paths, the player/strategy ends the partnership immediately. Because the deviator can end the partnership unilaterally, ending the partnership is the maximal punishment, i.e., the optimal penal code.

Let $\Omega = \cup_{t=1}^{\infty}(\{C, D\} \times \{C, D, \emptyset\})^{(t-1)}$ be the set of action profile paths in a partnership. Interpret that the first coordinate is the player's own action and the second coordinate is the current partner's 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 $t = 1, 2, \dots, \min\{|q|, |q'|\}$, $(q(1), \dots, q(t-1)) = (q'(1), \dots, q'(t-1)) \Rightarrow q_1(t) = q'_1(t)$;

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

The first condition requires that if observed path up to t is the same, the same own action is specified at t. This 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 $s(Q) \in S$ is a *simple strategy* if, in each period t,
(i) in the stage game, it plays according to the unique q_{1t} generated by Q and the observed path; and
(ii) in the continuation decision phase, it keeps the partnership if and only if the observed path is contained in Q.

We can extend the ordinary C-trigger strategy to our model as a simple strategy with the singleton set of acceptable path

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

In this case the strategy starts the game by playing C and ends the partnership as the punishment, as soon as deviation from (C, C) is observed. As noted after Lemma 1, this strategy is not a candidate of NSS and thus will not be discussed below. The myopic $\tilde{d}$-strategy is also a simple strategy with

$$Q_{\tilde{d}} = \{(D, \emptyset)\}.$$

In this section we consider *single-norm* simple strategy distributions, under which no voluntary separation occurs on the play path. A single-norm distribution may consist of multiple strategies, but they are in coordination so that their actions are mutually acceptable.

3.1. *Symmetric Action Simple Strategies*

In this subsection we focus on simple strategy distributions whose acceptable paths have only symmetric action profiles. In particular, we analyze *trust-building* strategies whose acceptable paths look like $\{(D, D), \ldots, (D, D), (C, C), \ldots\}$. Strategies which revert to (D, D) after some periods of (C, C) do not earn higher average payoffs than those that continue (C, C) forever after and thus are not candidates of NSD. To play (C, C) at some point, the partners need some periods of (D, D) due to Lemma 1.

DEFINITION. For any $T = 1, 2, 3, \ldots$, let c_T-strategy be the symmetric-action simple strategy with the singleton set of acceptable paths

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

We call the first T periods of c_T-strategy as the *trust-building periods* and the periods afterwards as the *cooperation periods*.

We identify a condition that strategies which differ from c_T in one-step (in particular during the cooperation periods) do not give a higher payoff than c_T-strategy.[4] Let p_T be the monomorphic strategy distribution consisting only of

[4]Note that the literal one-step deviation strategy is not feasible for our players since they cannot play differently across partnerships. However, if c_T-strategy in unimprovable by *infeasible* one-step deviation, then it is unimprovable within S.

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

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

The average payoff that a player with c_T-strategy expects to receive in the partnership with another c_T-strategy from t-th period on is denoted as $v^I(c_T, c_T, t)$ and is called *continuation average payoff in a match.* It is increasing for $t \leq T$, as less and less time is spent for trust-building, and it stays constant at c for any $t \geq T + 1$.

$$v^I(c_T, c_T, t) = \begin{cases} (1 - \delta^{2(T-t+1)})d + \delta^{2(T-t+1)}c & \text{if} \quad T \geq t \\ c & \text{if} \quad t \geq T+1 \end{cases}$$

Let $L(s, s', t)$ be the expected length of a match of s with s' from t-th period on; i.e., $L(s, s', t) = 1 + \delta^2 + \cdots + \delta^{2(T(s,s')-t)}$. When c_T is matched with another c_T, $L(c_T, c_T, t) = \frac{1}{1-\delta^2}$ for any t. If one follows c_T-strategy, the continuation average payoff is $v^I(c_T, c_T, t)$ for $L(c_T, c_T, t)$ periods, and for the rest ($L - L(c_T, c_T, t)$ periods), it is $v(c_T; p_T)$ since it has to go through the matching pool. By contrast, one-step deviation from c_T-strategy at t-th period (where $t > T$) gets the immediate payoff of g, but for the rest ($L - 1$ periods), the continuation average payoff is $v(c_T; p_T)$. Therefore, no one-step deviation during cooperation periods earns higher payoff than c_T-strategy if and only if

$$\begin{aligned} & g + (L-1)v(c_T; p_T) \leq L(c_T, c_T, t)v^I(c_T, c_T, t) + \{L - L(c_T, c_T, t)\}v(c_T; p_T) \\ \iff & g - v^I(c_T, c_T, t) \leq \{L(c_T, c_T, t) - 1\}[v^I(c_T, c_T, t) - v(c_T; p_T)]. \end{aligned}$$

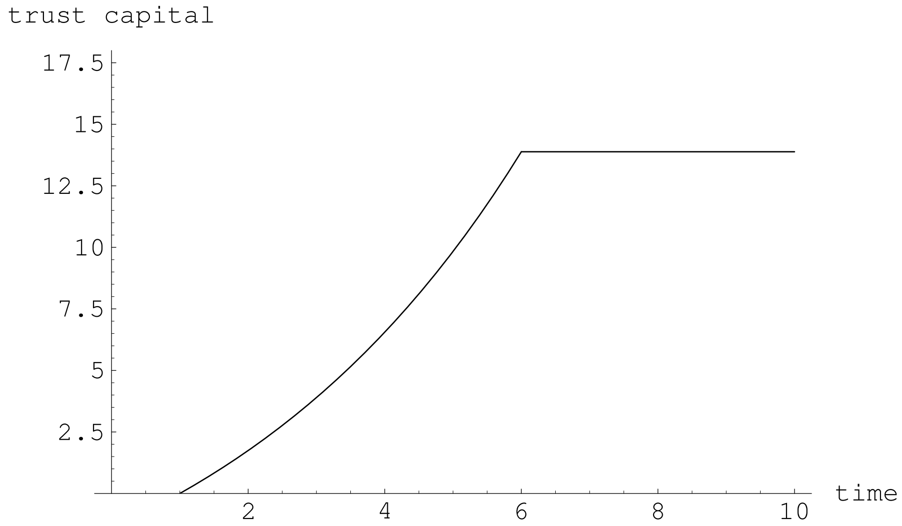

FIGURE 1. – The trust capital of c_T-strategy in a match with c_T.

(Parameter values: $c = 10, d = 1, \delta = 0.9, T = 5$.)

We call the RHS of the above inequality, $\{L(c_T, c_T, t) - 1\}[v^I(c_T, c_T, t) - v(c_T; p_T)]$, *trust capital* of c_T-strategy at the beginning of t-th period in a match with another c_T-strategy. It is initially zero and increases until the cooperation periods start.

While the average gain of deviation is $g - v^I(c_T, c_T, t)$, the loss of future average value is the trust capital. Since $L(c_T, c_T, t) = \frac{1}{1-\delta^2}$ and $v^I(c_T, c_T, t) = c$, the *best response condition* that c_T-strategy is better than any one-step deviation during the cooperation periods is

$$g - c \leq \frac{\delta^2}{1-\delta^2}[c - v(c_T; p_T)] \iff v(c_T; p_T) \leq \frac{c - (1-\delta^2)g}{\delta^2} =: v^{BR}. \qquad (5)$$

Since v^{BR} is independent of the length T of trust-building periods and $v(c_T; p_T)$ decreases as T increases, there is a lower bound to T that warrants (5).

Now we prove that in fact the best response 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 response condition can be written as

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

Given (G, T), define $\underline{\delta}_G(T)$ as the solution to

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

Then the best response condition (5) is satisfied if and only if $\delta \geq \underline{\delta}_G(T)$. It is easy to see that

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

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

$$\underline{\tau}(\delta, G) := argmin_{\tau \in \Re_{++}} \{\underline{\delta}_G(\tau) \mid \delta \geq \underline{\delta}_G(\tau)\}.$$

(See Figure 5 in Section 4.3.)

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

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., $T \geq \underline{\tau}(\delta, G)$. Strategies that differ from c_T-strategy off the play path do not give a higher payoff. Q.E.D.

Proposition 1 shows that for sufficiently high survival probabilities and sufficiently long trust-building periods, voluntarily repeated cooperation is sustained. 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 $\sqrt{\frac{g-c}{g-d}} = \underline{\delta}_G(\infty)$. This means that cooperation in VRPD requires more patience.

Next we investigate whether the monomorphic strategy distribution p_T, when it is a NE, is neutrally stable. In general, in order to check whether a Nash equilibrium strategy is a NSS, we only need to consider mutants that are best replies to the Nash equilibrium strategy.

LEMMA 4. *Suppose $p \in \mathcal{P}(\boldsymbol{S})$ is a NE. If a pure strategy $s' \in S$ invades p, 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 $T > \underline{\tau}(\delta, G)$, however, Lemma 3 (c) implies that such strategies are not alternative best reply. Therefore c_T-strategy is NSS for this case.

It may happen that (δ, G) allow an integer $\underline{\tau}(\delta, G)$. For this case, we consider an alternative best reply to p_T ($T = \underline{\tau}(\delta, G)$ hereafter) which earns the highest payoff when meeting itself. Among alternative best replies (that play D at some point in the cooperation periods), c_{T+1} earns the highest payoff when meeting itself. It suffices to identify a condition that c_{T+1}-strategy cannot invade p_T.

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

PROOF: See Appendix.

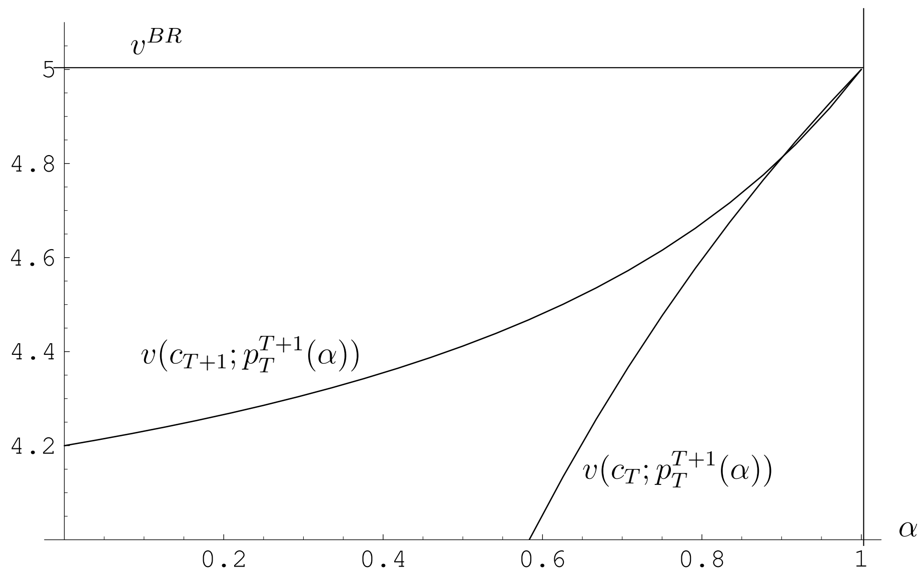

FIGURE 2. – The value functions of c_T-strategy and c_{T+1}-strategy when $T = \underline{\tau}(\delta, G)$.

(Parameter values: $g = 10, c = 6, d = 1, \ell = -1, \delta = \frac{2}{\sqrt{5}}, T = \underline{\tau}(\delta, G) = 1$.)

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

PROOF: See Appendix.

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 2.

LEMMA 7. *Take any G and any $\delta \in (\underline{\delta}_G(\infty), 1)$. Let $T = \underline{\tau}(\delta, G)$. 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 $\hat{\tau}(\delta, G)$ implicitly as the solution to

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

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

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 $T = \underline{\tau}(\delta, G)$. The RHS 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 the other 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. (See Figure 5 in Section 4.3.)

In sum, we have the following parametric condition for the existence of single-norm symmetric NSS.

PROPOSITION 2. *(a) For any G and any $\delta > \underline{\delta}_G(\infty)$ such that $\delta \neq \underline{\delta}_G(T)$ for any T, c_T-strategy is NSS for any $T \geq \underline{\tau}(\delta, G)$.*

(b) For any G and any $\delta > \underline{\delta}_G(\infty)$ such that $\delta = \underline{\delta}_G(T)$ for some T, c_T-strategy is NSS if and only if $T = \underline{\tau}(\delta, G) \geq \hat{\tau}(\delta, G)$.

In words, when δ is large enough, voluntarily repeated cooperation after trust-building is NSS with sufficient number of trust-building periods, even though myopic defection is not.

3.2. *Asymmetric Action Simple Strategies*

Next, we allow asymmetric action profiles in the acceptable paths. Again, if C is to be played on the play path, at least one period of (D, D) is required, and thus the equilibrium strategies have some trust-building periods. Let us focus on the one alternating (C, D) and (D, C) after the trust-building periods, which gives the most efficient and symmetric outcome when $2c < g + \ell$. (Other sequences can be analyzed similarly, and one can achieve the same payoffs by a symmetric strategy combinations, if the model is extended to allow role-dependent actions or correlated actions.)

Consider a two-strategy distribution consisting of the following simple strategies.

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}. Hence

$$\begin{aligned}
V^I(a_T,b_T) &= (1+\delta^2+\cdots+\delta^{2(T-1)})d+\delta^{2T}(1+\delta^4+\cdots)\ell+\delta^{2(T+1)}(1+\delta^4+\cdots)g \\
&= \frac{1-\delta^{2T}}{1-\delta^2}d+\frac{\delta^{2T}\ell+\delta^{2(T+1)}g}{1-\delta^4}, \\
V^I(b_T,a_T) &= (1+\delta^2+\cdots+\delta^{2(T-1)})d+\delta^{2T}(1+\delta^4+\cdots)g+\delta^{2(T+1)}(1+\delta^4+\cdots)\ell \\
&= \frac{1-\delta^{2T}}{1-\delta^2}d+\frac{\delta^{2T}g+\delta^{2(T+1)}\ell}{1-\delta^4}.
\end{aligned}$$

Let $p^b_{aT}(\alpha)$ be a two-strategy distribution with the support $\{a_T,b_T\}$ such that α of the population uses a_T and $(1-\alpha)$ uses b_T. Since any kind of match continues ad infinitum, $L(a_T,a_T)=L(a_T,b_T)=L(b_T,b_T)=\frac{1}{1-\delta^2}$.

$$\begin{aligned}
v(a_T;p^b_{aT}(\alpha)) &= \alpha\{(1-\delta^{2T})d+\delta^{2T}c\}+(1-\alpha)\{(1-\delta^{2T})d+\frac{\delta^{2T}\ell+\delta^{2(T+1)}g}{1+\delta^2}\} \\
v(b_T;p^b_{aT}(\alpha)) &= \alpha\{(1-\delta^{2T})d+\frac{\delta^{2T}g+\delta^{2(T+1)}\ell}{1+\delta^2}\} \\
&\quad +(1-\alpha)\{(1-\delta^{2(T+1)})d+\delta^{2(T+1)}c\}.
\end{aligned} \tag{8}$$

By computation

$$\begin{aligned}
&[v(a_T;p^b_{aT}(\alpha))-v(b_T;p^b_{aT}(\alpha))]\frac{1+\delta^2}{\delta^{2T}} \\
=\ &\alpha(1+\delta^2)\{d-\ell-(g-c)+\delta^2(c-d)\}-\{(d-\ell)-\delta^2(g-c)+\delta^4(c-d)\}.
\end{aligned}$$

Note that

$$\lim_{\delta\to 1}\{d-\ell-(g-c)+\delta^2(c-d)\}=2c-(g+\ell).$$

Therefore, $2c<g+\ell$ if and only if $v(a_T;p^b_{aT}(\alpha))$ crosses $v(b_T;p^b_{aT}(\alpha))$ from the above as α increases, i.e., there exists $\bar{\alpha}(\delta)$ such that

$$\alpha \gtreqless \bar{\alpha}(\delta) \iff v(b_T;p^b_{aT}(\alpha)) \gtreqless v(a_T;p^b_{aT}(\alpha)),$$

so that when $\alpha>\bar{\alpha}(\delta)$ (resp. $\alpha<\bar{\alpha}(\delta)$, b_T-strategy does better than a_T-strategy and thus α decreases (resp. increases). By computation,

$$\bar{\alpha}(\delta)=\frac{(d-\ell)-\delta^2(g-c)+\delta^4(c-d)}{(1+\delta^2)\{d-\ell-(g-c)+\delta^2(c-d)\}}.$$

Since $\lim_{\delta\to 1}\bar{\alpha}(\delta) = \frac{1}{2}$, for sufficiently large δ, $\bar{\alpha}(\delta) \in (0,1)$ is warranted.

By a similar logic to the one for the symmetric strategy profile, we can prove the best response condition that the alternating action distribution is NE (and NSD) for sufficiently long trust-building.

PROPOSITION 3. *Assume $g + \ell > 2c$. For sufficiently large δ, there exists $\tilde{T}(\delta)$ such that for any $T \geq \tilde{T}$, the two-strategy distribution $p^b_{aT}(\bar{\alpha}(\delta))$ is a NSD.*

PROOF: Similar to Propositions 1 and 2 and is omitted.

4. MULTI-NORM SIMPLE STRATEGY DISTRIBUTIONS

The literature on voluntarily repeated games has concentrated on single-norm, monomorphic equilibria so that no voluntary break-up occurs, except for sorting out inherent defectors under incomplete information case. (See concluding remarks 5.4.) We now investigate equilibria under which voluntary break-ups occur on the play path. There is no reason to believe that such equilibria do not exist or are less efficient than single-norm equilibria. Recall that our model is of complete information and with homogeneous players. Hence equilibrium break-up is interpreted as society having multiple norm, i.e., the acceptable paths of the players are not in coordination.

Below we deal with only symmetric-action simple-strategy distributions, i.e., c_T-strategies with different trust-building periods T, but a similar analysis can be done for asymmetric-action simple-strategy profiles.

4.1. *Bimorphic Distribution*

We investigate conditions that the two-strategy distribution (called *bimorphic* distribution) of $p_T^{T+1}(\alpha) = \alpha p_T + (1-\alpha)p_{T+1}$ is a NSD for some $\alpha_T^{T+1} \in (0,1)$. For a bimorphic distribution $p_T^{T+1}(\alpha_T^{T+1})$ to be a NSD, the following two conditions must hold.

Payoff Equalization and Within Distribution Stability:

$$\alpha \gtreqless \alpha_T^{T+1} \iff v(c_{T+1}; p_T^{T+1}(\alpha)) \gtreqless v(c_T; p_T^{T+1}(\alpha)). \tag{9}$$

If the above is satisfied, at α_T^{T+1}, c_T-strategy and c_{T+1}-strategy earn the same average payoff, and when $\alpha > \alpha_T^{T+1}$ (resp. $\alpha < \alpha_T^{T+1}$), c_{T+1}-strategy fares better than c_T-strategy so that α decreases (resp. c_T-strategy fares better so that α increases). Hence the bimorphic distribution $p_T^{T+1}(\alpha_T^{T+1})$ cannot be invaded by strategies that have the same play path as c_T or c_{T+1} under the bimorphic distribution.

Best Response Condition: Let $\hat{v}_T^{T+1} := v(c_T; p_T^{T+1}(\alpha_T^{T+1})) = v(c_{T+1}; p_T^{T+1}(\alpha_T^{T+1}))$. For any $t \geq T+2$,

$$g + (L-1)\hat{v}_T^{T+1} \leq L(c_T, c_T, t)v^I(c_T, c_T, t) + \{L - L(c_T, c_T, t)\}\hat{v}_T^{T+1}. \tag{10}$$

This condition implies that no one-step deviation during the cooperation periods of c_{T+1}-strategy can earn average payoff higher than c_T or c_{T+1} strategy. Since c_T and c_{T+1}-strategy have the same average payoff under $p_T^{T+1}(\alpha_T^{T+1})$, the RHS of (10) is the continuation average payoff of either strategy.

LEMMA 8. *For any $s \in \boldsymbol{S}$ such that s imitates c_T or c_{T+1} during their trust-building periods and chooses D at some point in their cooperation periods, $v(s; p_T^{T+1}(\alpha_T^{T+1})) \leq \hat{v}_T^{T+1}$ if and only if (10) holds.*

PROOF: Similar to Lemma 3 (c) and is thus omitted.

By computation, we can rewrite Best Response Condition as

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

Recall that by definition of $\underline{\tau}(\delta, G)$ (abbreviated as $\underline{\tau}$ below),

$$v(c_{\underline{\tau}}; p_{\underline{\tau}}^{\underline{\tau}+1}(1)) = v(c_{\underline{\tau}+1}; p_{\underline{\tau}}^{\underline{\tau}+1}(1)),$$

i.e., the best response condition is satisfied with equality at $\alpha = 1$ when the trust-building periods is exacatly $\underline{\tau}(\delta, G)$. Hence there is no bimorphic NSD with the support $\{c_{\underline{\tau}}, c_{\underline{\tau}+1}\}$. Figure 3 shows that, given (δ, G), as T *slightly* decreases below $\underline{\tau}(\delta, G)$, both $v(c_T; p_T^{T+1}(\alpha))$ and $v(c_{T+1}; p_T^{T+1}(\alpha))$ increases (uniformly for all $\alpha \in [0, 1]$) but they 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 $T < \hat{\tau}(\delta, G)$, using the same logic as Lemma 7. Hence there is a lower bound to δ to warrant $\hat{\tau}(\delta, G) \geq \underline{\tau}(\delta, G)$. (See Figure 5 in Section 4.3.)

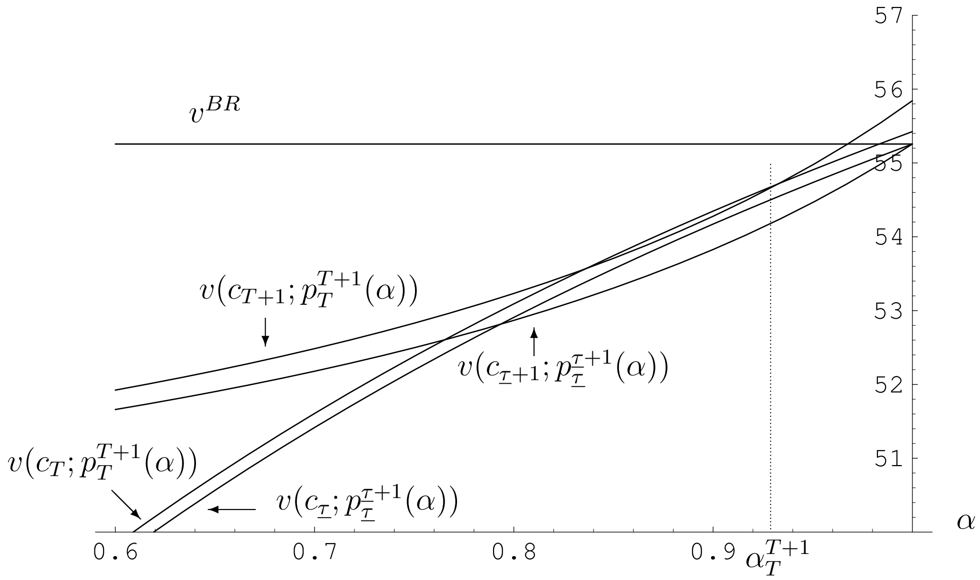

FIGURE 3. – The existence of a bimorphic NSD as T is slightly below $\underline{\tau}(\delta, G)$. (Parameter values: $g = 100, c = 61, d = 21, \ell = 20, \delta = 0.921556, T = 1.05, \underline{\tau}(\delta, G) = 1.1$.)

For any (δ, G, T), let

$$\begin{aligned}\Phi &= (1-\delta^2)(g-d) - \delta^2\{1-\delta^{2(T+1)}\}(c-d) - [(1-\delta^2)(d-\ell) \\ &\qquad +\{1-\delta^{2(T+1)}\}(c-d)][1-\delta^{2(T+1)}\}^2 > 0, \\ \Psi &= [(1-\delta^2)(d-\ell) + \{1-\delta^{2(T+1)}\}(c-d)][\{1-\delta^{2(T+1)}\}^2(c-d) - (1-\delta^2)(g-d)].\end{aligned}$$

PROPOSITION 4. *For any (δ, G) such that $\hat{\tau}(\delta, G) \geq \underline{\tau}(\delta, G)$,*

$$\frac{dv}{dT}\Big|_{T=\underline{\tau}(\delta,G)} = 2\delta^{2T}(\log\delta)[\frac{(c-d)\Phi+\Psi}{\Phi}].$$

Hence, if $(c-d)\Phi + \Psi < 0$, then (since $\log\delta < 0$) for T sufficiently close to $\underline{\tau}(\delta, G)$ but less than that, there exists $\alpha_T^{T+1} \in (0,1)$ such that $p_T^{T+1}(\alpha_T^{T+1})$ is a bimorphic NSD.

PROOF: See Appendix.

4.2. *Staggered Distribution: Finite Support*

We can extend the sufficient condition of the bimorphic NSDs for a general multi-norm NSDs with the support $\{c_T, c_{T+1}, \ldots, c_{T+K}\}$. Let $p_T^{T+K}(\alpha, \beta_1, \ldots, \beta_{K-1}) = \alpha p_T + (1-\alpha)\beta_1 p_{T+1} + (1-\alpha)(1-\beta_1)\beta_2 p_{T+2} + \cdots + (1-\alpha) \times_{k=1}^{K-1} (1-\beta_k) p_{T+K}$ be the strategy distribution.

The Payoff Equalizing condition is derived backwards: Given the fractions of $(\alpha, \beta_1, \ldots, \beta_{K-2})$, find $\beta_{K-1}^* \in (0,1)$ such that

$$\beta_{K-1} \gtreqless \beta_{K-1}^* \iff v(c_{T+K}; p_T^{T+K}(\beta_{K-1})) \gtreqless v(c_{T+K-1}; p_T^{T+K}(\beta_{K-1})). \tag{11}$$

Since c_{T+K-1} and c_{T+K} behave the same way against any c_{T+k}-strategy such that $k \leq K-2$, their payoff difference comes from only the matches with c_{T+K-1} and c_{T+K}. Let $\bar{V} := \sum_{k=0}^{K-2} L(c_{T+k}, c_{T+K-1}) v^I(c_{T+K-1}, c_{T+k}) = \sum_{k=0}^{K-2} L(c_{T+k}, c_{T+K}) v^I(c_{T+K}, c_{T+k})$, and $\gamma := (1-\alpha) \times_{k=1}^{K-2} (1-\beta_k)$. Then

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

Therefore the payoff equalizing and stable β_{K-1}^* is one of the two solutions to

$$v(c_{T+K-1}; p_T^{T+K}(\beta_{K-1})) = v(c_{T+K}; p_T^{T+K}(\beta_{K-1}))$$

which is a quadratic equation of β_{K-1}. Hence a similar analysis to the existence of α_T^{T+1} in 4.1 can be applied to the existence of β^*_{K-1}. Note that β^*_{K-1} is a function of $(\alpha, \beta_1, \ldots, \beta_{K-2})$.

Given $\beta^*_{K-1}(\alpha, \beta_1, \ldots, \beta_{K-2})$, and $(\alpha, \beta_1, \ldots, \beta_{K-3})$, let $\beta^*_{K-2} \in (0,1)$ be such that

$$\beta_{K-2} \gtreqless \beta^*_{K-2} \iff v(c_{T+K-1}; p_T^{T+K}(\beta_{K-2})) \gtreqless v(c_{T+K-2}; p_T^{T+K}(\beta_{K-2})). \quad (12)$$

And the process goes on until we reach $\alpha_T^{T+K} \in (0,1)$ such that

$$\alpha \gtreqless \alpha_T^{T+K} \iff v(c_{T+1}; p_T^{T+K}(\alpha)) \gtreqless v(c_T; p_T^{T+K}(\alpha)). \quad (13)$$

Since there are exploiters for c_{T+1}-strategy under p_T^{T+K}, $v(c_{T+1}; p_T^{T+K}(\alpha)) < v(c_{T+1}; p_T^{T+1}(\alpha))$. Therefore it is more difficult for the intersection α_T^{T+K} to exist than the bimorphic case. However, if α_T^{T+1} and α_T^{T+K} exist for some $K > 1$, then all inbetween Payoff Equalizing solutions α_T^{T+k} $(k = 1, 2, \ldots, K-1)$ exist. See Figure 4.

Once the Payoff Equalizing and Stable $(\alpha_T^{T+K}, \beta_1^*(\alpha_T^{T+K}), \ldots, \beta^*_{K-1}(\alpha_T^{T+K}))$ exists we can write the Best Response Condition similar to the bimorphic case.

Best Response Condition: Let $\hat{v}_T^{T+K} := v(c_T; p_T^{T+K}(\alpha_T^{T+K})) = v(c_{T+k}; p_T^{T+K}(\alpha_T^{T+K}))$ for all $k = 1, 2, \ldots, K$. For any $t \geq T + K + 1$,

$$g + (L-1)\hat{v}_T^{T+K} \leq L(c_T, c_T, t)v^I(c_T, c_T, t) + \{L - L(c_T, c_T, t)\}\hat{v}_T^{T+K}. \quad (14)$$

Using the same logic as in 4.1, we can identify at least a local condition near $\underline{\tau}(\delta, G)$ for the existence of K-morphic NSD.

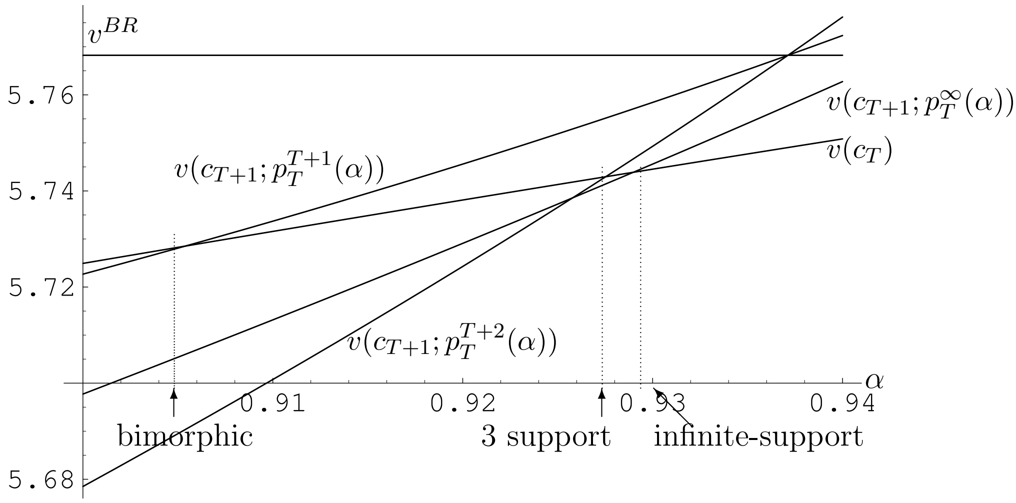

FIGURE 4. – Existence of Multi-Norm NSDs.

4.3. *Staggered Distribution: Infinite Support*

Finally we consider simple strategy distributions with the support $\{c_T, c_{T+1}, \ldots\}$ (for some T), 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 9. *Take any G and $T < \infty$. Let p be a stationary strategy distribution with the support $\{c_T, c_{T+1}, \ldots\}$. If $v(c_T; p) = v(c_{T+\tau}; p)$ for all $\tau = 1, 2, \ldots$, then the fraction of $c_{T+\tau}$-strategy is $\alpha(1-\alpha)^\tau$ for each $\tau = 0, 1, 2, \ldots$.*

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 logic is easily understood by the following Table II(a)-(c) showing the matching probability and the sequence of payoffs for c_T, c_{T+1}, and c_{T+2} within a match against c_T, c_{T+1}, and so on.

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	$\boldsymbol{d}$	$\cdots$	$\boldsymbol{d}$	$\boldsymbol{d}$	$\boldsymbol{d}$	$\boldsymbol{\ell}$	

Notice that the bold-faced sub-table of Table II(b) is identical to the Table II(a). This is because from the second period on, c_{T+1}-strategy behaves the same way as c_T-strategy against itself and against longer trust-building strategies. The conditional probabilities of meeting itself and longer trust-building strategies are also the same. 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.

LEMMA 10. *For any G and $T < \infty$,*

$$v(c_T; p_T^\infty(\alpha)) = v(c_{T+1}; p_T^\infty(\alpha)) \Rightarrow v(c_{T+\tau}; p_T^\infty(\alpha)) = v(c_T; p_T^\infty(\alpha)) \ \forall \tau = 1, 2, \ldots,$$

PROOF: (Can be omitted.) See Appendix.

Consider the following class of G:

$$\mathcal{G} = \{G \mid \exists T \in \{1, 2, \ldots\} \text{ such that } 1 \geq \underline{\delta}_G(T) > \hat{\delta}_G(T)\}$$

To see if such G exists, note that for any G, $\hat{\delta}_G(T)$ is increasing and $\underline{\delta}_G(T)$ is decreasing in T. For $T = 0$, $\hat{\delta}_G(0) = \sqrt{1 - \frac{c-d}{g-\ell}} < \infty = \underline{\delta}_G(0)$, and $\underline{\delta}_G(\infty) = \sqrt{\frac{g-c}{g-d}} < 1 = \hat{\delta}_G(\infty)$. Therefore the graphs of $\hat{\delta}_G(T)$ and $\underline{\delta}_G(T)$, when the time scale is extended to real numbers, have an intersection. Then some adjustment of the payoffs will give rise to $G \in \mathcal{G}$. (See Figure 5.)

LEMMA 11. *Take any $G \in \mathcal{G}$. For any $T \in \{1, 2, \ldots\}$ such that $\hat{\delta}_G(T) < \underline{\delta}_G(T)$, there exists $\delta^*_G(T) \in (\hat{\delta}_G(T), \underline{\delta}_G(T))$ such that for any $\delta \in [\delta^*_G(T), \underline{\delta}_G(T))$, there exists $\alpha^*(\delta)$ that satisfies:*

(a) $v(c_T; p^\infty_T(\alpha^*)) = v(c_{T+1}; p^\infty_T(\alpha^*))$, and

(b) $\Delta v^\infty_T(\alpha, \delta)$ is negatively sloping at α^*, where $\Delta v^\infty_T(\alpha, \delta) := v(c_T; p^\infty_T(\alpha)) - v(c_{T+1}; p^\infty_T(\alpha))$.

PROOF: See Appendix.

Property (a), together with Lemma 10, implies that all strategies in the support of $p^\infty_T(\alpha^*)$ have the same payoff. Property (b) implies that, near α^*, c_T-strategy does better than c_{T+1}-strategy if and only if $\alpha < \alpha^*$. Hence c_T or c_{T+1} cannot invade (increase the fraction in) $p^\infty_T(\alpha^*)$. Moreover, if $c_{T+\tau}$ ($\tau \geq 2$) increases the fraction in $p^\infty_T(\alpha^*)$, it is as if c_{T+1} increases (i.e., α decreases). Then c_T gets higher payoff, and $c_{T+\tau}$ ($\tau \geq 2$) cannot invade in the sense of increasing the fraction.

It remains to identify the sufficient conditions of δ and T such that no other strategy earns higher payoff, i.e., an incentive constraint. By a similar logic to the monomorphic NSD, it suffices to consider strategies that differ on the play path after cooperation periods have started. For each $\tau = 0, 1, 2, \ldots$, let s_τ be the strategy that imitates $c_{T+\tau}$-strategy for the first $T + \tau + 1$ periods (that is, to build trust for $T + \tau$ periods and then play C once, so that it is clear that the partner is $c_{T+\tau}$-strategy if the partnership continues) and then plays D in $T+\tau+2$. A sufficient condition for such strategy to be unable to invade $p_T^\infty(\alpha^*)$ is

$$v(c_{T+\tau}; p_T^\infty(\alpha^*)) > v(s_\tau; p_T^\infty(\alpha^*)). \tag{15}$$

Note that, among on-path deviations during the cooperation periods, s_τ strategy earns the highest payoff, due to the discounting.

LEMMA 12. *For any* (G, T)*, there exists* $\delta_G^{**}(T) < \underline{\delta}_G(T)$ *such that for any* $\delta \in (\delta_G^{**}(T), \underline{\delta}_G(T))$,

$$v(c_{T+\tau}; p_T^\infty(\alpha^*(\delta))) > v(s_\tau; p_T^\infty(\alpha^*(\delta)))$$

for any $\tau = 0, 1, 2, \ldots$.

PROOF. See Appendix.

PROPOSITION 5. *Take any* $G \in \mathcal{G}$. *For any* T *such that* $\hat{\delta}_G(T) < \underline{\delta}_G(T)$ *and for any* $\delta \in (\max\{\delta_G^*(T), \delta_G^{**}(T)\}, \underline{\delta}_G(T))$*, there is a neutrally stable polymorphic strategy distribution of the form* $p_T^\infty(\alpha^*(\delta))$ *for some* $\alpha^*(\delta) \in (0, 1)$.

PROOF: (Can be omitted.) From lemma 11 and lemma 12, it suffices to prove that other strategies that differ on the play path from $\{c_T, c_{T+1}, \ldots\}$ during the trust-building periods or that end a partnership with the same strategy do not earn higher payoff, which is shown in Appendix. Q.E.D.

Therefore, cooperation and exploitation can co-exist. Figure 5 below summarizes the relationship between the survival rate δ and the sufficient trust-building periods of symmetric single-norm NSDs and multi-norm NSDs.

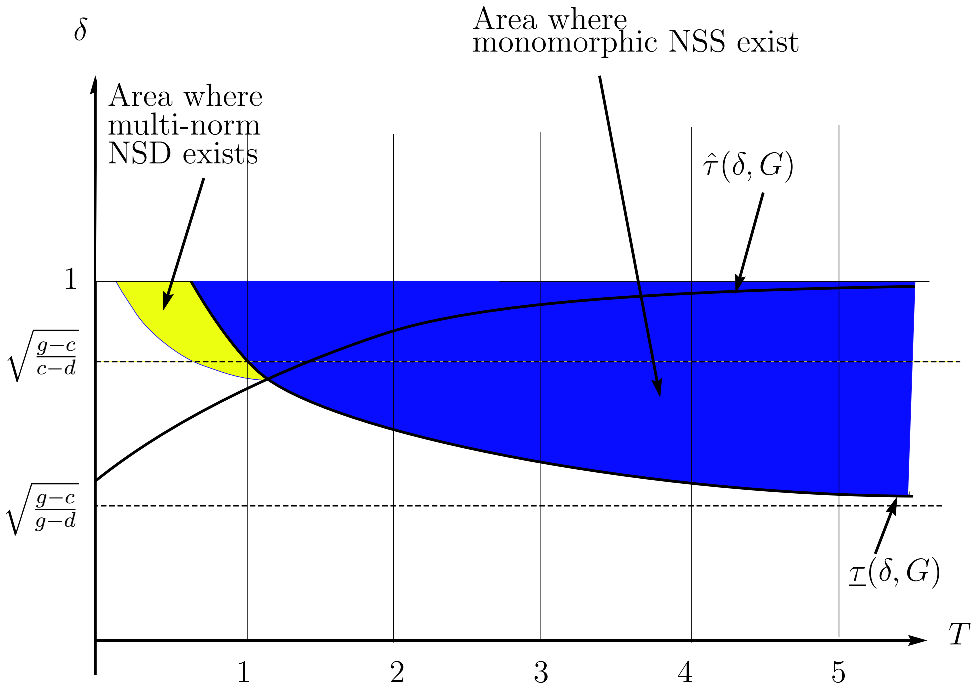

FIGURE 5. – Parametric summary of NSDs.

4.4. *Efficiency*

Finally, let us compare the efficiency between single-norm NSDs and multi-norm NSDs.

As we have seen, the shortest trust-building period among monomorphic NSS is longer than the shortest trust-building period of multi-norm NSDs. (See Figure 5 above.)

Recall that the average payoff of each strategy in the support of a polymorphic NSD is (for any K finite or infinity)

$$\begin{aligned} v(c_T; p_T^{T+K}(\alpha_T^{T+K}(\delta))) &= \alpha_T^{T+K}(\delta)\frac{L(c_T, c_T)}{L(c_T; p_T^{T+K}(\alpha_T^{T+K}(\delta)))}v^I(c_T, c_T) \\ &\quad +(1-\alpha_T^{T+K}(\delta))\frac{L(c_T, c_{T+1})}{L(c_T; p_T^{T+K}(\alpha_T^{T+K}(\delta)))}v^I(c_T, c_{T+1}) \\ &= v(c_{T+k}; p_T^{T+K}(\alpha^*(\delta))) \qquad \forall k = 1, 2, \ldots, \infty \end{aligned}$$

For δ sufficiently close to but less than $\underline{\delta}_G(T)$, $\alpha_T^{T+K}(\delta) \approx 1$ and thus

$v(c_T; p_T^{T+K}(\alpha_T^{T+K}(\delta))) \approx v(c_T; p_T) > v(c_{T+1}; p_{T+1})$ where the latter is the shortest monomorphic NSS's average payoff.

Therefore, it is possible that a multi-norm NSD is more efficient than any single-norm monomorphic NSS. This is due to the randomly matched partner's exploitation as a disciplining device. If you are lucky to start the cooperation periods with the current partner, it is not beneficial to betray and go back to the random matching pool, in which there are diverse players with longer trust-building periods.

It is quite interesting that diversity can mean early cooperation as well as higher average payoff than single-norm NSDs.

5. CONCLUDING REMARKS

5.1. *Efficiency Wage and Three Types of Sanctions*

Our model describes a society where players meet with a stranger to play a voluntarily repeated 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 non-cooperation 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 different ways by which trust is generated;

(a) With new partner, player must spend a certain lengths of trust-building (TB) periods, playing one-shot NE. This is the case when c_T-strategy constitutes a monomorphic NSD with $T \geq 1$.

(b) With a positive probability, new partner will play a different length of TB periods. Partnership with such stranger is a mismatch, forcing the partnership to break up prematurely. This is the case when, e.g., c_0- and c_1-strategies coexist in a staggered polymorphic NSD.

There is an additional mechanism which creates trust if we allow matching probability to be less than one.

(c) 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 monomorphic 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 $u \in (0, 1)$ she 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", 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 incentive constraint for c_T-strategy to be followed during the cooperation periods is

$$\begin{aligned} & g - v^I(c_T, c_T, t) \leq \{L(c_T, c_T, t) - 1\}\{v^I(c_T, c_T, t) - v^0(c_T; p_T)\} \\ \iff \quad & g - c \leq \frac{\delta^2}{1 - \delta^2}\{c - v^0(c_T; p_T)\}. \end{aligned}$$

Clearly, if (5) is satisfied, the above constraint is also satisfied. Moreover, the BR constraint can be satisfied even for c_0 for sufficiently large u, and cooperation without trust building period becomes a self-sustaining state.

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 current 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. *Cheap Talk*

Recall that, under favorable environments, c_1-strategy can invade the population of $\tilde{d}$-strategy as an equilibrium entrant. c_1-strategy proposes to keep the partnership even after (D, D), and 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 shall provide a rough sketch of what would happen if we allow cheap talk with neoloigism at the beginning of each matching.

Assume that when two players are newly matched, they simultaneously choose and send a message $m \in M$ from a countable set M to her 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 VRPD with cheap talk consists of (m, σ) such that:

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

In the following, however, in order to ease the notation we shall denote pure strategy by σ, omitting the message m she sends to her partner. We chose such convention because we focus on the following two types of strategies; babbling strategy where message choice has no meaninful contents and neologism strategy where the message is "neologism".

In what follows, $\boldsymbol{S}^{CT}$ is the set of all pure strategies of VRPD with cheap talk, which is the extension of $\boldsymbol{S}$ defined for the original VRPE without cheap talk.

1. "Babbling" strategy: $s \in \boldsymbol{S}$ is extended as a degenerate strategy $\sigma^B \in \boldsymbol{S}^{CT}$ which uses a constant-vaued function $\sigma^B(m) = s$ for all $m \in M$. This strategy makes initial message exchange meaningless because $s \in \boldsymbol{S}$ is played regardless of the message received from the partner.

2. "Neologism" strategy: Different VRPD strategies $s, s' \in \boldsymbol{S}$ are played depending upon the message receieved from the partner. E.g., suppose the current population consists of a babbling strategy $\sigma^B \in \boldsymbol{S}^{CT}$ where $\sigma^B(m) = s \in \boldsymbol{S}$ for any $m \in M$. Against this monomorphic strategy distribution, consider an entrant population who uses a strategy $\sigma^N \in \boldsymbol{S}^{CT}$ such that

(a) it announces a neologism message, i.e., a message which is not used by the current population, and

(b) $\sigma^N(m) = s$ when m is not the neologism, while

(c) $\sigma^N(m') = s' \neq s$ when m' is the neologism.

With this neologism strategy, σ^N, 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.

For each pure strategy $s \in \boldsymbol{S}$, we write corresponding babbling strategy of the cheap talk model (actually, "set of strategies" because message choice is arbitrary) as $\sigma^B(s) \in \mathcal{P}(\boldsymbol{S}^{CT})$. Similarly, we can extend a strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$ of the ordinary VRPD to an associated babbling strategy distribution of the cheap talk game. For the ease notations, we shall write this distribution as $\sigma^B(p) \in \mathcal{P}(\boldsymbol{S}^{CT})$, with the superscript B.

Given an incumbent babbling strategy distribution, $\sigma^B(p') \in \mathcal{P}(\boldsymbol{S})$, consider a neologism strategy which plays $s \in \boldsymbol{S}$ if and only if both partners use the neologism message. Again in order to ease notation, we shall denote such strategy (i.e., strategy which tries to enter p' using the neologism-contingent play of s) as $\sigma^N(s;p')$.

As is well-known, babbling extension of a Nash equilibrium of the original model is always a Nash equilibrium of the cheap-talk model because message exchange does not alter players' incentives if all players use babbling strategy.

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

PROOF: Obvious.

On the other hand, some NSDs of non-cheap talk model are invaded by a neologism strategy as an equilibrium entrant in the cheap talk model. However, in the cheap-talk model, it is possible to signal to start the cooperation periods earlier and earlier to eventually violate the Best Response condition. Hence we need to require that entrants must be a best response to the post-entry distribution, to avoid the non-existence.

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

In Section 4, we showed that, given G, either the single-norm NSD with the minimum trust-building periods or one of the multi-norm NSD is the most effcient (i.e., whose expected payoff provides the largest value) NSD of VRPD. Let $p^* \in \mathcal{P}(\boldsymbol{S})$ of VRPD be this most efficient (whose expected payoff is the highest) NSD of VRPD. Let $\sigma^B(p^*) \in \mathcal{P}(S^{CT})$ be the associated babbling strategy distribution.

Clearly, no NSD provides higher average payoff than the most efficient NSD, and with cheap talk no strategy can invade the most efficient NSD as an equlibrium entrant. Thus, we have the following result.

Proposition 6. *$\sigma^B(p^*) \in \mathcal{P}(S^{CT})$ is a NSDEE with cheap talk.*

Proof: Obvious.

5.3. *Drift and Limit of Solution Concept*

In this paper, we have used Nash Equilibrium (NE), Neutrally Stable Distribution (NSD), and Neutrally Stable Distribution against Equilibrium Entrants (NSDEE) as our solution concept. Because component game of our (random matching) model is VRPD which is an extensive form game, however, any strategy distribution leaves many unreaced nodes. Hence, the concept of NSD is not sufficiently restrictive to identify reasonably restricted set of strategy distributions.

Limitation of the cocept, 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 p_1 is the most efficient NSD. Being a NSDEE, once entier society starts to use c_T (or $\sigma^B(c_T) if cheap talk is allowed$), 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 $\tilde{c}_1 \in \boldsymbol{S}$:

$t = 1$: 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 $\tilde{c}_1$. Thus, strarting from p_1, strategy distribution may drift to any distribution $\beta p_1 + (1-\beta)\tilde{p}_1$ with $\beta \in [0, 1]$, where $\tilde{p}_1$ is the monomorphic distribution consisting only of $\tilde{c}_1$.

However, $\tilde{c}_1$ 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 $\tilde{c}_1$. Note that c_∞ can receive average payoff of only d in the strategy distribution p_1 and its payoff is strictly lower than c_1. However if drifts make β sufficiently large, c_∞ starts to drive out c_1. Eventually, strategy distribution may become p_∞, the monomorphic 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 $\tilde{p}_1$, from $\tilde{p}_1$ 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.

5.4. *Related Literature and Future Research Agenda*

Several papers have previously analyzed the voluntarily repeated prisoner's dilemma, though not as fully as this paper does. These literature has pointed out two factors that facilitate cooperation under the VRPD type games.

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

Datta (1996) and Kranton (1996a) consider a complete information, two-player, voluntarily repeated game similar to ours. The stage game is a continuum-action prisoner's dilemma, representing borrower-lender or gift exchange situations. Therefore the players in their model can gradually increase the "level of cooperation", which makes the same disciplining system as our c_T-strategy. Based upon non-evolutionary model-setting, Kranton (1996a) emphasizes that if partners "renegotiate", they would want to start cooperation immediately, which leads to non-existence of equilibrium. She then shifts to incomplete information by introducing "defective" type whose discount factor is zero so that initial low level of cooperation is rational to sort out these types.

Ghosh and Ray (1996) also consider a similar incomplete information model to Kranton's. In a related work, Watson (2002) shows that partners would choose to gradually increase the stakes of the relationship.

Carmichael and MacLeod (1997) formulate a complete information (i.e., all players have identical discount factor) evolutionary model with initial gift exchange stage added to the voluntarily repeated prisoner's dilemma. Mutual gift exchange, which incurs a positive cost to givers but provides no value to receivers, works the same way as the "level of cooperation" adjustment.

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

Second, "heterogeneity" may help cooperation. With incomplete information model Rob and Yang (2005), independently written from ours, shows that repeated cooperation from the outset of a partnership can be an equilibrium among heterogeneous players. In their model, there are three types of players; bad type who always plays D, good type who always plays C, and rationanl type who tries to maximize their payoff.

The logic is as follows. Existence of bad type players makes it valuable to (1) keep and cooperate with either good or rational type partners, and (2) to find out bad type partners as soon as possible. Thus, a rational player should cooperate from the beginning to be distinguished from the bad-type.

Our result is much starker than Rob and Yang. Our model does not rely on heterogeous "type" and incomplete information. Instead, bad (longer trust-buidling) strategy emerges endogenously as a polymorphic NSD. We also elucidate that there may be an equliibrium strategy distribution (NSD) with more than two (possible infinite) 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

$$v(c_T; p_T) - v(s; p_T) \geq 0 \iff \delta^2\frac{1-\delta^{2T}}{1-\delta^2}(c-d) \geq g-c.$$

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, then for any $s \in \mathit{supp}(p)$,

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

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

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

for sufficiently small $\epsilon > 0$. By letting $\epsilon \to 0$, we obtain

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

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

PROOF OF LEMMA 5:

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

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

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

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:

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

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 c_{T+1} can invade p_T (thus the first bracket is postive) 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 (**??**) and (**??**) become

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

By differentiation,

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

By computation,

$$\begin{aligned}\mu_T'(\alpha) &= \frac{L(c_T, c_T)L(c_T, c_{T+1})}{[\alpha L(c_T, c_T) + (1-\alpha)L(c_T, c_{T+1}))]^2}\\ &\to \frac{L(c_T, c_{T+1}))}{L(c_T, c_T)} = 1 - \delta^{2(T+1)} \quad \text{as} \quad \alpha \to 1,\\ \mu_{T+1}'(\alpha) &= \frac{L(c_{T+1}, c_T)L(c_T, c_{T+1})}{[\alpha L(c_{T+1}, c_T) + (1-\alpha)L(c_{T+1}, c_{T+1}))]^2}\\ &\to \frac{L(c_{T+1}, c_{T+1})}{L(c_{T+1}, c_T)} = \frac{L(c_T, c_T)}{L(c_T, c_{T+1})} = \frac{1}{1-\delta^{2(T+1)}}, \quad \text{as} \quad \alpha \to 1.\end{aligned}$$

At $\delta = \delta_G(T)$,

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

Therefore, at $\delta = \delta_G(T)$,

$$\begin{aligned}\frac{\partial \Delta \tilde{v}_T}{\partial \alpha}(1, \delta_G(T)) &= \frac{L(c_T, c_{T+1}))}{L(c_T, c_T)}\{v^I(c_T, c_T) - v^I(c_T, c_{T+1})\}\\ &\quad - \frac{L(c_T, c_T)}{L(c_T, c_{T+1})}\{v^I(c_{T+1}, c_T) - v^I(c_{T+1}, c_{T+1})\},\\ &= (1-\delta^{2(T+1)})\frac{\delta^{2T}(1-\delta^2)(g-\ell)}{1-\delta^{2(T+1)}} - \frac{1}{1-\delta^{2(T+1)}}\delta^{2T}(1-\delta^2)(c-d)\\ &= \delta^{2T}(1-\delta^2)\Big\{(g-\ell) - \frac{c-d}{1-\delta^{2(T+1)}}\Big\}.\end{aligned}$$

Q.E.D.

PROOF OF PROPOSITION 4: For the ease of computations, we introduce following notations.

(a) $\underline{\tau}(\delta, G) = T_0$

(b) $f(T) = v(c_T, c_T)$

(c) $h(T) = v(c_T, c_{T+1})$

(d) $g(T) = v(c_{T+1}, c_T)$

(e) $k(T) = v(c_{T+1}, c_{T+1})$

(f) $j(T) = v(c_{T+1}, c_{T+2})$

(g) $F(T,\alpha) = v(c_T; p_T^{T+1}(\alpha))$
$= m(T,\alpha)v(c_T, c_T) + [1 - m(T,\alpha)]v(c_T, c_{T+1})$
$= v(c_T; p_T^{\infty}(\alpha))$

(h) $G(T,\alpha) = v(c_{T+1}; p_T^{T+1}(\alpha))$
$= n(T,\alpha)v(c_{T+1}, c_T) + [1 - n(T,\alpha)]v(c_{T+1}, c_{T+1})$

(i) $H(T,\alpha) = v(c_{T+1}; p_T^{\infty}(\alpha))$
$= x(T,\alpha)v(c_{T+1}, c_T) + y(T,\alpha)v(c_{T+1}, c_{T+1}) + [1 - x(T,\alpha) - y(c_{T+1}, c_{T+2})]v(c_{T+1}, c_{T+2})$

(j) $m(T,\alpha) = \frac{\alpha L(c_T, c_T)}{\alpha L(c_T, c_T) + (1-\alpha)L(c_T, c_{T+1})} = \frac{\alpha p(T)}{\alpha p(T) + (1-\alpha)q(T)}$

(k) $n(T,\alpha) = \frac{\alpha L(c_{T+1}, c_T)}{\alpha L(c_{T+1}, c_T) + (1-\alpha)L(c_{T+1}, c_{T+1})} = \frac{\alpha q(T))}{\alpha q(T) + (1-\alpha)p(T)}$

(l) $x(T,\alpha) = \frac{\alpha L(c_{T+1}, c_T)}{\alpha L(c_{T+1}, c_T) + \alpha(1-\alpha)L(C_{T+1}, c_{T+1}) + (1-\alpha^2)L(c_{T+1}, c_{T+2})}$
$= \frac{\alpha q(T)}{\alpha q(T) + \alpha(1-\alpha)p(T) + (1-\alpha^2)r(T)}$

(m) $y(T,\alpha) = \frac{\alpha(1-\alpha)L(c_{T+1}, c_{T+1})}{\alpha L(c_{T+1}, c_T) + \alpha(1-\alpha)L(C_{T+1}, c_{T+1}) + (1-\alpha^2)L(c_{T+1}, c_{T+2})}$
$= \frac{\alpha(1-\alpha)p(T)}{\alpha q(T) + \alpha(1-\alpha)p(T) + (1-\alpha^2)r(T)}$

(n) $p(T) = L(c_T, c_T) = L(c_{T+1}, c_{T+1})$

(o) $q(T) = L(c_T, c_{T+1}) = L(c_{T+1}, c_T) = L(c_T, c_{T+2})$

(p) $r(T) = L(c_{T+1}, c_{T+2})$

(q) $v^{IC} = c - \frac{1-\delta^2}{\delta^2}(g - c)$

By definition,

$$F(T_0, 1) = f(T_0) = G(T_0, 1) = g(T_0) := v_0.$$

Moreover, note that v^{IC} is independent of T and:

$$\frac{\partial v^{IC}}{\partial T} = 0. \tag{18}$$

We now analyze the effect of a small change (decrease) of T, which is denoted as a change from T_0 to $T = T_0 + dT$. If we denote **payoff equalizing** strategy proportion by $\alpha_T^{T+1}(T)$, and the associated **equalized payoff** by $v(T)$, the following identity must hold:

$$F(T, \alpha_T^{T+1}(T)) = G(T, \alpha_T^{T+1}(T)) = v(T).$$

Differentiating this identity:

$$\frac{\partial F}{\partial T} + \frac{\partial F}{\partial \alpha}\frac{d\alpha_T^{T+1}(T)}{dT} = \frac{\partial G}{\partial T} + \frac{\partial G}{\partial \alpha}\frac{d\alpha_T^{T+1}(T)}{dT} = \frac{dv}{dT}. \tag{19}$$

Rearranging,

$$\frac{d\alpha_T^{T+1}(T)}{dT} = \frac{\frac{\partial F}{\partial T} - \frac{\partial G}{\partial T}}{\frac{\partial G}{\partial \alpha} - \frac{\partial F}{\partial \alpha}}. \tag{20}$$

Note, by assumption,

$$\frac{\partial F}{\partial T}\Big|_{(T,\alpha)=(T_0,1)} - \frac{\partial G}{\partial T}\Big|_{(T,\alpha)=(T_0,1)} > 0$$

because $T < \underline{\tau}(\delta, G)$ and as T becomes small, $g(T) - f(T)$ increases. On the other hand,

$$\frac{\partial G}{\partial \alpha}\Big|_{(T,\alpha)=(T_0,1)} > \frac{\partial F}{\partial \alpha}\Big|_{(T,\alpha)=(T_0,1)}$$

because $T < \hat{\tau}(\delta, G)$. It follows that, in the neighbourhood of $(T, \alpha) = (T_0, 1)$,

$$\frac{d\alpha_T^{T+1}(T)}{dT} > 0.$$

In view of (18), if:

$$dv = \left\{ \frac{\partial F(T, \alpha_T^{T+1}(T))}{\partial T} + \frac{\partial F(T, \alpha_T^{T+1}(T))}{\partial \alpha} \frac{d\alpha_T^{T+1}(T)}{dT} \right\}\Big|_{T=T_0} \times dT < 0 \text{ when } dT < 0, \quad (21)$$

or **if the sign of (19) is positive, then bimorphic distribution $p_T^{T+1}(\alpha(T))$ is NE for T which is sufficiently close to T_0.**

Note that $\frac{\partial F}{\partial \alpha} > 0$ but $\frac{\partial F}{\partial T} < 0$ and the sign of (19) is *a priori* ambiguous.

Differentiation of $F(T, \alpha)$

Because:

$$F(T, \alpha) = m(T, \alpha) f(T) + [1 - m(T, \alpha)] h(T),$$

$$\frac{\partial F}{\partial T} = \frac{\partial m}{\partial T}[f(T) - h(T)] + m(T, \alpha) f'(T) + [1 - m(T, \alpha)] h'(T), \text{ and} \quad (22)$$

$$\frac{\partial F}{\partial \alpha} = \frac{\partial m}{\partial \alpha}[f(T) - h(T)]. \quad (23)$$

Note:

$$\frac{\partial m(T, \alpha)}{\partial T} = \frac{\alpha p'(T)[\alpha p(T) + (1-\alpha) q(T)] - \alpha p(T)[\alpha p'(T) + (1-\alpha) q'(T)]}{[\alpha p(T) + (1-\alpha) q(T)]^2}$$

$$= \frac{\alpha(1-\alpha)[p'(T) r(T) + q'(T) p(T)]}{[\alpha p(T) + (1-\alpha) q(T)]^2} \to 0 \text{ as } \alpha \to 1, \quad (24)$$

$$\frac{\partial m(T, \alpha)}{\partial \alpha} = \frac{p(T)[\alpha p(T) + (1-\alpha) q(T)] - [p(T) - q(T)] \alpha p(T)}{[\alpha p(T) + (1-\alpha) q(T)]^2}$$

$$= \frac{p(T) q(T)}{[\alpha p(T) + (1-\alpha) q(T)]^2} \to \frac{q(T)}{p(T)} \text{ as } \alpha \to 1, \quad (25)$$

$$m(T, \alpha) = \frac{\alpha p(T)}{\alpha p(T) + (1-\alpha) q(T)} \to 1 \text{ as } \alpha \to 1. \quad (26)$$

Differentiation of $G(T, \alpha)$

Because:

$$G(T, \alpha) = n(T, \alpha) g(T) + [1 - n(T, \alpha)] k(T),$$

$$\frac{\partial G}{\partial T} = \frac{\partial n(T,\alpha)}{\partial T}[g(T)-k(T)] + n(T,\alpha)g'(T) + [1-n(T,\alpha)]k'(T), \quad (27)$$

$$\frac{\partial G}{\partial \alpha} = \frac{\partial n(T,\alpha)}{\partial \alpha}[g(T)-k(T)]. \quad (28)$$

Note:

$$\begin{aligned}
\frac{\partial n(T,\alpha)}{\partial T} &= \frac{\alpha q'(T)[\alpha q(T)+(1-\alpha)p(T)] - [\alpha q'(T)+(1-\alpha)p'(T)]\alpha q(T)}{[\alpha q(T)+(1-\alpha)p(T)]^2} \\
&= \frac{\alpha(1-\alpha)[q'(T)p(T)-q(T)p'(T)]}{[\alpha q(T)+(1-\alpha)p(T)]^2} \to 0 \text{ as } \alpha \to 1, \quad (29) \\
\frac{\partial n(T,\alpha)}{\partial \alpha} &= \frac{q(T)[\alpha q(T)+(1-\alpha)p(T)] - \alpha q(T)[q(T)-p(T)]}{[\alpha q(T)+(1-\alpha)p(T)]^2} \\
&= \frac{q(T)p(T)}{[\alpha q(T)+(1-\alpha)p(T)]^2} \to \frac{p(T)}{q(T)} \text{ as } \alpha \to 1, \quad (30) \\
n(T,\alpha) &= \frac{\alpha\alpha q(T)}{\alpha q(T)+(1-\alpha)p(T)} \to 1 \text{ as } \alpha \to 1. \quad (31)
\end{aligned}$$

Computation of $\alpha'(T_0)$

In view of (24), (25), (26), (29), (30) and (31):

$$\begin{aligned}
\frac{\partial F}{\partial T}\Big|_{\alpha=1} &= f'(T), \quad (32) \\
\frac{\partial F}{\partial \alpha}\Big|_{\alpha=1} &= \frac{q(T)}{p(T)}[f(T)-h(T)], \quad (33) \\
\frac{\partial G}{\partial T}\Big|_{\alpha=1} &= g'(T), \quad (34) \\
\frac{\partial G}{\partial \alpha}\Big|_{\alpha=1} &= \frac{p(T)}{q(T)}[g(T)-k(T)]. \quad (35)
\end{aligned}$$

Because $\frac{d\delta^T}{dT} = \delta^T(\log\delta)$, straightforward computations yield:

$$\begin{aligned}
f(T) &= v(c_T, c_T) = (1-\delta^{2T})d + \delta^{2T}c, \\
f'(T) &= 2\delta^{2T}(\log\delta)(c-d), \\
h(T) &= v(c_T, c_{T+1}) = \frac{(1-\delta^{2T})d + \delta^{2T}(1-\delta^2)\ell}{1-\delta^{2(T+1)}}, \\
f(T)-h(T) &= (1-\delta^{2T})d + \delta^{2T}c - \frac{(1-\delta^{2T})d + \delta^{2T}(1-\delta^2)\ell}{1-\delta^{2(T+1)}} \\
&= \frac{(1-\delta^{2T})[1-\delta^{2(T+1)}-1]d + \delta^{2T}[(1-\delta^{2(T+1)})c - (1-\delta^2)\ell]}{1-\delta^{2(T+1)}} \\
&= \frac{\delta^{2T}}{1-\delta^{2(T+1)}}[(1-\delta^2)(c-\ell) + \delta^2(1-\delta^{2T})(c-d)], \\
p(T) &= L(c_T, c_T) = \frac{1}{1-\delta^2}, \\
q(T) &= L(c_T, c_{T+1}) = \frac{1-\delta^{2(T+1)}}{1-\delta^2}.
\end{aligned}$$

Substituting these values into (32) and (33):

$$\begin{aligned}
\frac{\partial F}{\partial T}\Big|_{\alpha=1} &= f'(T) = 2\delta^{2T}(\log\delta)(c-d), && (36)\\
\frac{\partial F}{\partial \alpha}\Big|_{\alpha=1} &= \frac{q(T)}{p(T)}[f(T)-h(T)]\\
&= [1-\delta^{2(T+1)}]\frac{\delta^{2T}[(1-\delta^2)(c-\ell)+\delta^2(1-\delta^{2T})(c-d)]}{1-\delta^{2(T+1)}}\\
&= \delta^{2T}[(1-\delta^2)(c-\ell)+\delta^2(1-\delta^{2T})(c-d)]\\
&= \delta^{2T}[(1-\delta^2)(d-\ell)+(1-\delta^{2(T+1)})(c-d)]. && (37)
\end{aligned}$$

Similarly:

$$\begin{aligned}
g(T) &= v(c_{T+1},c_T) = \frac{(1-\delta^{2T})d+\delta^{2T}(1-\delta^2)g}{1-\delta^{2(T+1)}}\\
g'(T) &= \frac{[-2\delta^{2T}d+2\delta^{2T}(1-\delta^2)g](\log\delta)[1-\delta^{2(T+1)}]+2\delta^{2(T+1)}(\log\delta)[(1-\delta^{2T})d+\delta^{2T}(1-\delta^2)g]}{[1-\delta^{2(T+1)}]^2}\\
&= \frac{2\delta^{2T}(\log\delta)}{[1-\delta^{2(T+1)}]^2}[(1-\delta^2)(1-\delta^{2(T+1)})g-(1-\delta^{2(T+1)})d+\delta^2(1-\delta^{2T})d+\delta^{2(T+1)}(1-\delta^2)g]\\
&= \frac{2\delta^{2T}(\log\delta)}{[1-\delta^{2(T+1)}]^2}[(1-\delta^2)g-(1-\delta^2)d]\\
&= \frac{2\delta^{2T}(\log\delta)}{[1-\delta^{2(T+1)}]^2}(1-\delta^2)(g-d),\\
g(T)-k(T) &= v(c_{T+1},c_T)-v(c_{T+1},c_{T+1})\\
&= \frac{(1-\delta^{2T})d+\delta^{2T}(1-\delta^2)g}{1-\delta^{2(T+1)}}-(1-\delta^{2(T+1)})d-\delta^{2(T+1)}c\\
&= \frac{(1-\delta^{2T})d+\delta^{2T}(1-\delta^2)g-[1-\delta^{2T}+\delta^{2T}(1-\delta^2)](1-\delta^{2(T+1)})d-(1-\delta^{2(T+1)})\delta^{2(T+1)}c}{1-\delta^{2(T+1)}}\\
&= \frac{\delta^{2T}(1-\delta^2)g-\delta^{2T}(1-\delta^2)d+(1-\delta^{2(T+1)})\delta^{2(T+1)}d-(1-\delta^{2(T+1)})\delta^{2(T+1)}c}{1-\delta^{2(T+1)}}\\
&= \frac{\delta^{2T}}{1-\delta^{2(T+1)}}[(1-\delta^2)(g-d)-\delta^2(1-\delta^{2(T+1)})(c-d)],\\
q(T) &= L(c_{T+1},c_T) = \frac{1-\delta^{2(T+1)}}{1-\delta^2},\\
p(T) &= L(c_{T+1},c_{T+1}) = \frac{1}{1-\delta^2}.
\end{aligned}$$

Substituting these values into (34) and (35):

$$\begin{aligned}
\frac{\partial G}{\partial T}\Big|_{\alpha=1} &= g'(T) = \frac{2\delta^{2T}(1-\delta^2)(\log\delta)}{[1-\delta^{2(T+1)}]^2}(g-d), && (38)\\
\frac{\partial G}{\partial \alpha}\Big|_{\alpha=1} &= \frac{p(T)}{q(T)}[g(T)-k(T)]\\
&= \frac{\delta^{2T}}{[1-\delta^{2(T+1)}]^2}[(1-\delta^2)(g-d)-\delta^2(1-\delta^{2(T+1)})(c-d)]. && (39)
\end{aligned}$$

Hence:

$$\begin{aligned}
&\frac{\partial F}{\partial T}\Big|_{\alpha=1} - \frac{\partial G}{\partial T}\Big|_{\alpha=1} \\
&= 2\delta^{2T}(\log\delta)\big\{c - d - \frac{(1-\delta^2)(g-d)}{[1-\delta^{2(T+1)}]^2}\big\} \\
&= 2\delta^{2T}(\log\delta)\frac{[\Lambda(T)]^2(c-d) - (1-\delta^2)(g-d)}{[\Lambda(T)]^2}, \text{ and}
\end{aligned} \tag{40}$$

$$\begin{aligned}
&\frac{\partial G}{\partial \alpha}\Big|_{\alpha=1} - \frac{\partial F}{\partial \alpha}\Big|_{\alpha=1} \\
&= \delta^{2T}\Big\{\frac{(1-\delta^2)(g-d) - \delta^2(1-\delta^{2(T+1)})(c-d)}{[1-\delta^{2(T+1)}]^2} - [(1-\delta^2)(d-\ell) + (1-\delta^{2(T+1)})(c-d)]\Big\} \\
&= \frac{\delta^{2T}}{[\Lambda(T)]^2}\{(1-\delta^2)(g-d) - \delta^2\Lambda(T)(c-d) - [(1-\delta^2)(d-\ell) + \Lambda(T)(c-d)][\Lambda(T)]^2\} > 0,
\end{aligned} \tag{41}$$

where:

$$\Lambda(T) = 1 - \delta^{2(T+1)}.$$

Substituting (40) and (41) into (20):

$$\begin{aligned}
\frac{d\alpha_T^{T+1}(T)}{dT}\Big|_{T=T_0} &= \frac{\frac{\partial F}{\partial T}\big|_{\alpha=1} - \frac{\partial G}{\partial T}\big|_{\alpha=1}}{\frac{\partial G}{\partial \alpha}\big|_{\alpha=1} - \frac{\partial F}{\partial \alpha}\big|_{\alpha=1}} \\
&= \frac{2(\log\delta)\{[\Lambda(T)]^2(c-d) - (1-\delta^2)(g-d)\}}{(1-\delta^2)(g-d) - \delta^2\Lambda(T)(c-d) - [(1-\delta^2)(d-\ell) + \Lambda(T)(c-d)][\Lambda(T)]^2}.
\end{aligned} \tag{42}$$

Computation of $\frac{dv}{dT}$

By (37) and (42):

$$\begin{aligned}
&\frac{\partial F(T, \alpha_T^{T+1}(T))}{\partial \alpha} \times \frac{d\alpha_T^{T+1}(T)}{dT}\Big|_{T=T_0} \\
&= \frac{\delta^{2T}[(1-\delta^2)(d-\ell) + \Lambda(T)(c-d)]2(\log\delta)\{[\Lambda(T)]^2(c-d) - (1-\delta^2)(g-d)\}}{(1-\delta^2)(g-d) - \delta^2\Lambda(T)(c-d) - [(1-\delta^2)(d-\ell) + \Lambda(T)(c-d)][\Lambda(T)]^2} \\
&:= 2\delta^{2T}(\log\delta)\frac{\Psi}{\Phi},
\end{aligned} \tag{43}$$

where:

$$\begin{aligned}
\Psi &:= [(1-\delta^2)(d-\ell) + \Lambda(T)(c-d)]\{[\Lambda(T)]^2(c-d) - (1-\delta^2)(g-d)\}, \\
\Phi &:= (1-\delta^2)(g-d) - \delta^2\Lambda(T)(c-d) - [(1-\delta^2)(d-\ell) + \Lambda(T)(c-d)][\Lambda(T)]^2.
\end{aligned}$$

Substituting (36) and (43) into (19):

$$\begin{aligned}
\frac{dv}{dT}\Big|_{T=T_0} &= \frac{\partial F(T, \alpha_T^{T+1}(T))}{\partial T} + \frac{\partial F(T, \alpha_T^{T+1}(T))}{\partial \alpha} \times \frac{d\alpha_T^{T+1}(T)}{dT}\Big|_{T=T_0} \\
&= 2\delta^{2T}(\log\delta)\big[c - d + \frac{\Psi}{\Phi}\big] \\
&= 2\delta^{2T}(\log\delta)\big[\frac{(c-d)\Phi + \Psi}{\Phi}\big].
\end{aligned} \tag{44}$$

Because $\log\delta < 0$, it follows then:

$$\frac{dv}{dT} = 2\delta^{2T}(\log\delta)\left\{\frac{(c-d)\Phi+\Psi}{\Phi}\right\} > 0 \iff (c-d)\Phi+\Psi < 0.$$

If the latter is satisfied, the bimorphic distribution $p_T^{T+1}(\alpha_T^{T+1})$ is a NSD if T is smaller than but sufficiently close to $T_0 = \underline{\tau}(\delta, G)$. Q.E.D.

PROOF OF LEMMA 9: 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

$$V(c_t;p,t+1) = \alpha_t\{\frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2}V(c_t;p)\} + (1-\alpha_t)\{\ell + \delta V(c_t;p)\}. \tag{45}$$

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

$$\begin{aligned} V(c_{t+1};p,t+1) &= \alpha_t\{g+\delta V(c_{t+1};p)\} \\ &\quad +(1-\alpha_t)\{d+\delta(1-\delta)V(c_{t+1};p)+\delta^2 V(c_{t+1};p,t+2)\}. \end{aligned} \tag{46}$$

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 (46) becomes

$$\begin{aligned} V(c_{t+1};p,t+1) &= \alpha_t\{g+\delta V(c_{t+1};p)\} \\ &\quad +(1-\alpha_t)\{d+\delta(1-\delta)V(c_{t+1};p)+\delta^2 V(c_t;p,t+1)\} \\ \iff V(c_{t+1};p,t+1) &= \frac{1}{1-(1-\alpha_t)\delta^2}\big[\alpha_t\{g+\delta V(c_{t+1};p)\} \\ &\quad +(1-\alpha_t)\{d+\delta(1-\delta)V(c_{t+1};p)\}\big]. \end{aligned} \tag{47}$$

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

$$V(c_t;p) = V(c_{t+1};p). \tag{48}$$

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

$$V(c_t;p,t+1) = V(c_{t+1};p,t+1). \tag{49}$$

(49) implies that the RHS of (45) and (47) must be the same. Using (48) and letting $V^*(p) = V(c_t;p) = V(c_{t+1};p)$, α_t must satisfy

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

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 10: From Table II(a), the long-run payoff of c_T-strategy is decomposed as

$$\begin{aligned} V(c_T; p_T^\infty(\alpha)) &= \alpha V(c_T, c_T; p_T^\infty(\alpha)) \\ &\quad +(1-\alpha)V(c_T, c_{T+1}; p_T^\infty(\alpha)). \end{aligned} \tag{50}$$

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

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

where the last equality uses (50).

Equivalently we can write the above as

$$\begin{aligned} [1-(1-\alpha)\delta(1-\delta)]V(c_{T+1}; p_T^\infty(\alpha)) &= \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) + (1-\alpha)d \\ &\quad +(1-\alpha)\delta^2 V(c_T; p_T^\infty(\alpha)). \end{aligned} \tag{51}$$

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

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

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

$$\begin{aligned} [1-(1-\alpha)\delta(1-\delta)]V(c_{T+2}; p_T^\infty(\alpha)) &= \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) + (1-\alpha)d \\ &\quad +(1-\alpha)\delta^2 V(c_{T+1}; p_T^\infty(\alpha)). \end{aligned} \tag{52}$$

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 (51) and (52) 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 LEMMA 11: We prove this lemma by a series of steps.

(a) *For any* (G, T, α, δ), $v(c_T; p_T^\infty(\alpha)) = v(c_T; p_T^{T+1}(\alpha))$.

Proof of (a): Against c_T-strategy, all strategies with longer trust-building behave the same way.

(b) *For any* (G, T, α, δ), *let* $\Delta v_T^\infty(\alpha, \delta) := v(c_T; p_T^\infty(\alpha)) - v(c_{T+1}; p_T^\infty(\alpha))$. *Then for any* (G, T, δ), $\Delta v_T^\infty(1, \delta) = \Delta v_T(1, \delta)$.

Proof of (b): Clearly when $\alpha = 1$, the value differences between c_T and c_{T+1} under p_T^∞ and p_T^{T+1} are the same.

(c) *For any* (G, T, δ), $\Delta v_T^\infty(0, \delta) < 0$.

Proof of (c): By computation,

$$\begin{aligned}
\Delta v_T^\infty(0,\delta) &= v^I(c_T, c_{T+1}) - v^I(c_{T+1}, c_{T+2}) \\
&= \frac{(1-\delta^{2T})d + \delta^{2T}(1-\delta^2)\ell}{1-\delta^{2(T+1)}} - \frac{(1-\delta^{2(T+1)})d + \delta^{2(T+1)}(1-\delta^2)\ell}{1-\delta^{2(T+2)}} \\
&= \frac{1}{(1-\delta^{2(T+1)})(1-\delta^{2(T+2)})}\Big[\{(1-\delta^{2T})d + \delta^{2T}(1-\delta^2)\ell\}(1-\delta^{2(T+2)}) \\
&\qquad -\{(1-\delta^{2(T+1)})d + \delta^{2(T+1)}(1-\delta^2)\ell\}(1-\delta^{2(T+1)})\Big]
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{(1-\delta^{2(T+1)})(1-\delta^{2(T+2)})}\Big[-(d-\ell)\{1-\delta^{2(T+1)} - (1-\delta^{2(T+2)})(1-\delta^{2T})\}\Big] \\
&< \frac{1}{(1-\delta^{2(T+1)})(1-\delta^{2(T+2)})}\Big[-(d-\ell)\{1-\delta^{2(T+1)} - (1-\delta^{2(T+1)})(1-\delta^{2T})\}\Big] \\
&= \frac{1}{(1-\delta^{2(T+1)})(1-\delta^{2(T+2)})}\Big[-(d-\ell)(1-\delta^{2(T+1)})\delta^{2T}\Big] < 0.
\end{aligned}$$

(d) *For any* (G, T, δ) *and any* $\alpha < 1$, $\Delta v_T^\infty(\alpha, \delta) > \Delta v_T(\alpha, \delta)$.

Proof of (d): Since c_{T+1} cannot be exploited under the two-strategy distribution p_T^{T+1} while it is exploited by strategies with longer trust-building periods under p_T^∞, $v(c_{T+1}; p_T^{T+1}(\alpha)) > v(c_{T+1}; p_T^\infty(\alpha))$. From (a), the statement holds.

Finally, we combine the above to prove the lemma. (b), (c), and (d) together imply that, for a given (G, T, δ), the graph of $\Delta v_T^\infty(\alpha, \delta)$ is uniformly above the graph of $\Delta v_T(\alpha, \delta)$ except at $\alpha = 1$ and both graph starts from a negative value at $\alpha = 0$. Hence, if there exists α such that $\Delta v_T(\alpha, \delta) = 0$ and $\frac{\partial \Delta v_T}{\partial \alpha}(\alpha, \delta) < 0$, then the desired α^* with the same properties for Δv_T^∞ also exists. (See Figure 2.)

The existence of such α for Δv_T is warranted if $\delta > \hat{\delta}_G(T)$ so that the slope of $\Delta \tilde{v}_T$ is negative, and if $\delta < \underline{\delta}_G(T)$ but sufficiently close to is so that $\Delta \tilde{v}_T(\alpha, \delta) > 0$ near $\alpha = 1$. Q.E.D.

PROOF OF LEMMA 12: Fix an arbitrary $\tau = 0, 1, 2, \ldots$. By computation,

$$\begin{aligned}
v^I(c_{T+\tau}, c_{T+\tau}) &= (1-\delta^{2(T+\tau)})d + \delta^{2(T+\tau)}c, \\
v^I(s_\tau, c_{T+\tau}) &= \frac{1}{1-\delta^{2(T+\tau+2)}}\Big[(1-\delta^{2(T+\tau)})d + \delta^{2(T+\tau)}(1-\delta^2)c + \delta^{2(T+\tau+1)}(1-\delta^2)g\Big].
\end{aligned}$$

For $\delta \approx \underline{\delta}_G(T)$,

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

Hence

$$\begin{aligned}
v^I(s_\tau, c_{T+\tau}) &\approx \frac{1}{1-\delta^{2(T+\tau+2)}}\Big[(1-\delta^{2(T+\tau)})d + \delta^{2(T+\tau)}(1-\delta^2)c \\
&\qquad +\delta^{2(T+\tau+1)}\{\delta^2(1-\delta^{2T})(c-d)+(1-\delta^2)c\}\Big] \\
&= \frac{1}{1-\delta^{2(T+\tau+2)}}\Big[(1-\delta^{2(T+\tau)})d + \delta^{2(T+\tau)}c - \delta^{2(T+\tau+1)}c \\
&\qquad +\delta^{2(T+\tau+1)}\{-\delta^2(1-\delta^{2T})d+(1-\delta^{2(T+1)})c\}\Big] \\
&= \frac{1}{1-\delta^{2(T+\tau+2)}}\Big[\Big\{(1-\delta^{2(T+\tau)})d + \delta^{2(T+\tau)}c\Big\}(1-\delta^{2(T+\tau+2)}+\delta^{2(T+\tau+2)}) \\
&\qquad -\delta^{2(T+\tau+2)}\{\delta^{2T}c+(1-\delta^{2T})d\}\Big] \\
&= v^I(c_{T+\tau}, c_{T+\tau}) - \frac{\delta^{2(T+\tau+2)}}{1-\delta^{2(T+\tau+2)}}\delta^{2T}(1-\delta^{2\tau})(c-d) < v^I(c_{T+\tau}, c_{T+\tau}).
\end{aligned}$$

Hence for δ sufficiently close to $\underline{\delta}_G(T)$, the in-match average payoff is smaller for s_τ. Moreover, it is easy to see that

$$r(c_{T+\tau}, c_{T+\tau}) = \frac{1}{1+\delta} > r(s_\tau, c_{T+\tau}) = \frac{1-\delta^{2(T+\tau+2)}}{1+\delta}.$$

By definition,

$$\begin{aligned}
v(c_{T+\tau}; p_T^\infty(\alpha)) &= \Big\{w + \alpha(1-\alpha)^\tau r(c_{T+\tau}, c_{T+\tau})v^I(c_{T+\tau}, c_{T+\tau}) \\
&\quad +(1-\alpha)^{\tau+1} r(c_{T+\tau}, c_{T+\tau+1})v^I(c_{T+\tau}, c_{T+\tau+1})\Big\} \\
&\quad /\Big\{R_1 + \alpha(1-\alpha)^\tau r(c_{T+\tau}, c_{T+\tau}) + (1-\alpha)^{\tau+1} r(c_{T+\tau}, c_{T+\tau+1})\Big\}
\end{aligned}$$

where $w = \sum_{k=0}^{\tau-1}\alpha(1-\alpha)^k r(c_{T+\tau}, c_{T+k})v^I(c_{T+\tau}, c_{T+k})$ and $R_1 = \sum_{k=0}^{\tau-1}\alpha(1-\alpha)^k r(c_{T+\tau}, c_{T+k})$.

Notice that s_τ behaves the same way as $c_{T+\tau}$ against c_{T+k} for $k = 0, 1, \ldots, \tau-1$ and $c_{T+\tau+1}$ and strategies with longer trust-building periods. Hence

$$\begin{aligned}
v(s_\tau; p_T^\infty(\alpha)) &= \Big\{w + \alpha(1-\alpha)^\tau r(s_\tau, c_{T+\tau})v^I(s_\tau, c_{T+\tau}) \\
&\quad +(1-\alpha)^{\tau+1} r(c_{T+\tau}, c_{T+\tau+1})v^I(c_{T+\tau}, c_{T+\tau+1})\Big\} \\
&\quad /\Big\{R_1 + \alpha(1-\alpha)^\tau r(s_\tau, c_{T+\tau}) + (1-\alpha)^{\tau+1} r(c_{T+\tau}, c_{T+\tau+1})\Big\}
\end{aligned}$$

Let $R_2 := \alpha(1-\alpha)^\tau r(c_{T+\tau}, c_{T+\tau})$ and $R_3 := (1-\alpha)^{\tau+1} r(c_{T+\tau}, c_{T+\tau+1})$. Then

$$\begin{aligned}
v(c_{T+\tau}; p_T^\infty(\alpha)) &= \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big)\frac{R_1+R_3}{R_1+R_2+R_3} + \frac{R_2 v^I(c_{T+\tau}, c_{T+\tau})}{R_1+R_2+R_3} \\
&= \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big) \\
&\quad + \frac{R_2}{R_1+R_2+R_3}\Big[v^I(c_{T+\tau}, c_{T+\tau}) \\
&\qquad - \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big)\Big]. \qquad (53)
\end{aligned}$$

Let $R_2' := \alpha(1-\alpha)^\tau r(s_\tau, c_{T+\tau})$. Then

$$\begin{aligned} v(s_\tau,;p_T^\infty(\alpha)) &= \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big)\frac{R_1+R_3}{R_1+R_2'+R_3} + \frac{R_2' v^I(s_\tau, c_{T+\tau})}{R_1+R_2'+R_3} \\ &= \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big) \\ &\quad + \frac{R_2'}{R_1+R_2'+R_3}\Big[v^I(s_\tau, c_{T+\tau}) \\ &\qquad -\Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big)\Big]. \end{aligned} \tag{54}$$

Since $R_2 > R_2'$, $\frac{R_1+R_3}{R_1+R_2+R_3} < \frac{R_1+R_3}{R_1+R_2'+R_3}$, which implies that $\frac{R_2'}{R_1+R_2'+R_3} < \frac{R_2}{R_1+R_2+R_3}$. Therefore the second term of (53) is larger than that of (54). Q.E.D.

PROOF OF PROPOSITION 5: We consider all on-path deviations that make a difference in the average payoff.

On-path deviations during the "common" trust-building periods $t = 1, 2, \ldots, T$: The on-path history is unique and of the form $\{(D, D), \ldots, (D, D)\}$. Possible deviation types that make difference in the payoffs are:

(a) Play e after on-path history, during the common TB.

Recall the logic of Lemma 2 (for monomorphic distribution). We showed that such strategy has average payoff d but any c_T strategy under p_T has average payoff more than d since it is a convex combination of c and d. Now we cannot use c_T but can use c_∞ which has the same payoff as c_T under α^*. c_∞ earns g after TB, against any c_τ where $\infty > \tau \geq T$. Hence the average payoff of c_∞ is more than d and thus it is better than choosing e during TB.

(b) Play C after on-path history, during the common TB.

Clearly, strategies in this class have smaller average payoff than d under $p_T^\infty(\alpha)$.

Thanks to the new definition that during TB, only (D, D) will induce k, we do not need to distinguish further devisions after C during TB.

On-path deviations in $t \geq T+1$: note that there are three kinds of on-path histories after the common trust-building periods.

1. $\{(D, D)^{t-1}\}$: This occurs when both partners had TB not less than $t-1$. For the continuation decison node, add one more (D, D).

   Action choice phase: Since both C and D are on-path actions we do not need to check.

   Continuation decision phase: The analysis is the same as (a) above.

2. $\{(D, D)^\tau, (C, C)^{t-\tau}\}$ for some $\tau \geq T$: This occurs when both partners had the same τ periods of TB. For the continuation decision node, add one more (C, C).

Action choice phase: The incentive constraint is proved to be satisfied in Lemma 12.

Continuation decision phase: If a strategy chooses C but e afterwards during the cooperation periods, the payoff is less than the above deviation strategy.

3. $\{(D, D)^{t-1}, (C, D)\}$: This is relevant only at the continuation decision node in t. This happens when one partner had $t-1$ periods of trust-building, while the other had a longer TB.

   However, by the definition of c_T strategy, the partner will choose e and thus your decision does not matter.

Q.E.D.

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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.