Voluntarily Repeated Prisoner's Dilemma*

by

Takako Fujiwara-Greve
Keio University

and

Masahiro Okuno-Fujiwara
University of Tokyo.

Preliminary and incomplete.
Final version will be available soon.

This draft: May 23, 2005

Abstract: We introduce a new model of voluntarily repeated games where the length of repetition is endogenously determined. We consider a large society of homogeneous players, in which players are randomly matched in each period to play prisoner's dilemma as well as to choose whether to play the game again with the same partner. When newly matched, players have no information about the past actions of each other. Therefore ordinary folk theorem does not apply. We show that neutrally stable strategy (NSS) must have some periods of trust-building followed by disciplinary cooperation (cooperate and keep the partnership if and only if the partner cooperated). Myopic defection is not NSS even though it is a Nash equilibrium strategy. The minimum length of trust-building depends on the gain of deviation and the death probability. When the number of trust-building periods is even smaller than this minimum, polymorphic NSS may exist, in which a variety of trust-building periods co-exist, and thus some are exploiting others even though the players are homogenous. When cheap talk is introduced at the beginning of the game, most efficient NSS becomes the unique NSS strategy.

Key words: voluntarily repeated prisoner's dilemma, evolution, cheap-talk.
JEL classification number: C 73 (Stochastic and dynamic games).

*We thank useful comments from Nobue Suzuki, Akihiko Matsui, Chiaki Hara, and seminar participants at KIER, Kyoto University and Keio University. All errors are ours.

1 Introduction

In this paper we introduce a model of voluntarily repeated games in a large society. Players are randomly matched to play a component game (specifically 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. While there is a large society that players belong to, each direct interaction (a partnership) is voluntarily separable. 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.

In our model, the incentive to free-ride is stronger than in ordinary repeated games, in which players are confined, for three reasons. First, a deviator can exit the partnership immediately after the deviation, escaping from any future punishment by the victim(s). Second, the information flow is severely limited. Actions within a partnership are assumed not to be revealed to players outside of the partnership, and, due to the random death, it is impossible to tell whether a player is in the matching pool because of a partnership separation or by birth. Third, the large society and random death make it impossible for a player who may know a deviator to start the contagion of defection (Kandori, 1992) effectively. Therefore, the known disciplining strategies such as trigger strategies (Fudenberg and Maskin, 1986) and contagion of defection do not sustain cooperation in our model. Our model describes a large anonymous society, which needs a different type of discipline from those of a society of directly interacting long-run players.

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. However, in the real-world, cooperative mode of behavior is still prevalent. We think that there is a disciplining system embedded in a society which supports cooperation even among boundedly rational players with little information about the past of each other.

Our idea is as follows. Voluntary termination of a partnership is not only useful for a deviator but also an option for the victim to escape from the potential future exploitation and at the same time to force the deviator into the random matching pool. If a newly made partnership can only produce much lower payoff, then returning to the matching pool serves as a punishment.[1] In order to create such payoff difference, one cannot start cooperating from the outset of a new partnership. Instead, there must be a *trust-building period*, in which partners do not cooperate but keep the partnership

[1]In the main body of the paper we assume that the probably of getting a new match in the pool is 1.

to build up *trust capital*, which will be lost once the current partnership ends. Then no player would deviate from the strategies with large enough trust capital.

We show that (1) playing cooperation from the first period of a match does not consitute even a Nash equilibrium in our model, (2) a myopic strategy which defects and ends a partnership immediately is a Nash equilibrium strategy but not neutrally stable, (3) sufficiently long period of trust-building with disciplinary cooperation afterwards is not only a Nash equilibrium but also neutrally stable strategy, and (4) in some cases, various trust-building periods co-exist in an equilibrium so that some are exploiting others. We also consider some extensions of the model such as allowing the probability of matching be less than 1 and allowing cheap talk before starting VRPD. When the matching probability is less than 1, it is easier to sustain cooperation. When cheap-talk is allowed, there is a unique EES strategy (Swinkels, 1992), which has the shortest equilibrium trust-building period.

This paper is organized as follows. In Section 2, we describe the basic model and stability concepts. In Section 3, we focus on equilibria in which players voluntarily keep the partnership and cooperate repeatedly. In Section 4, we consider matching frictions in relation to efficiency wage. In Sections 5, we consider drift and define another stability concept similar to EES (Swinkels, 1992) and, in Section 6, we analyze cheap-talk model with EES. Section 7 concludes the paper.

2 Model and Stability Concepts

2.1 Model

Consider a society with a continuum of players, each of whom may die in each period $t = 1, 2, \dots$ 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 simultanously 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. (See Figure 1 for the outline of the game.)

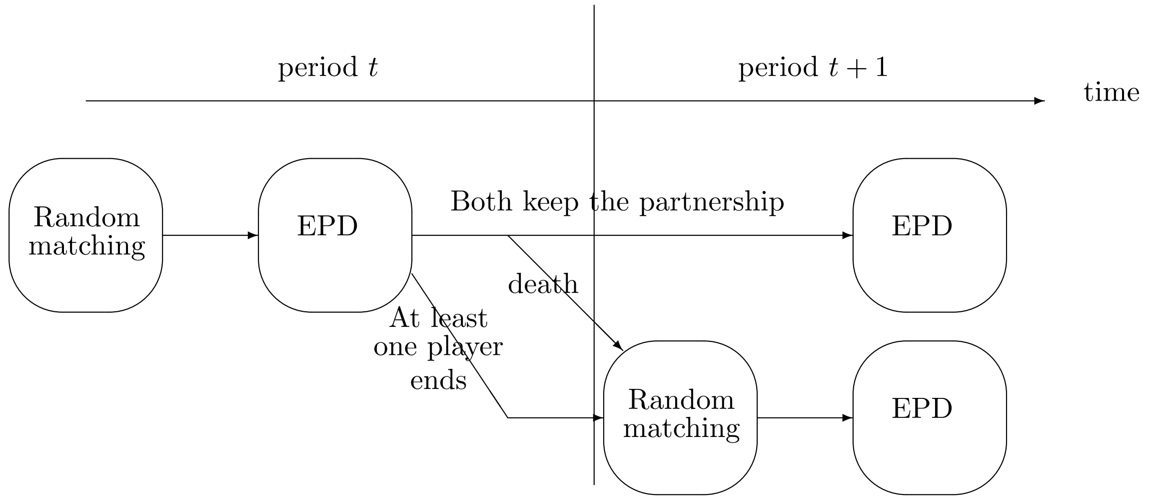

Figure 1: Outline of VRPD

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 $g>c>d>\ell$.[2]

P1 \ P2	C	D
C	c, c	ℓ, g
D	g, ℓ	d, d

Table 1: Payoff of PD

[2]We can normalize these payoffs such as $c=1$ and/or $d=0$. However, because later we introduce a payoff to be accrued when the player has no partner, we leave these payoffs as they are.

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. In t-th period of a match, let $(x_{1t}, x_{2t}) \in \{C, D\}^2$ be the action profile in PD, where the subscript 1 or 2 indicates the player role in this match (which is randomly assigned when a match is formed). Then a *partnership history* of VRPD at the beginning of $t = 2, 3, \ldots$, conditional upon the partnership being still alive is $h_t = \{(x_{1\tau}, x_{2\tau})\}_{\tau=1}^{t-1}$, while $h_0 = \{\emptyset\}$. Let

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

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

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

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, there is an associated stationary distribution of social states. See Appendix A for the details. 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.

Below we show how the lifetime payoffs are computed for a player in a match and for a player in the matching pool, waiting to be randomly matched.

When a strategy $s \in \boldsymbol{S}$ is matched with another strategy $s' \in \boldsymbol{S}$, the *expected number of days* s spends with s' will be denoted as $D(s,s')$ 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

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

When s is matched with s', the *portion of his lifetime* s expects to spend within the match with s' is:

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

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

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

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

Example 1. An extension of the trigger strategy in the ordinary (non-voluntary) repeated prisoner's dilemma is as follows:

$t = 1$: Play C and keep the partnership for any observation.

$t \geq 2$: Play C if the partnership history of PD actions consists only of (C,C). Otherwise play D. Keep the partnership for any observation.

If two players endowed with this trigger strategy tr are matched, then $D(tr,tr) = \frac{1}{1-\delta^2}$, $r(tr,tr) = \frac{1}{1+\delta}$, and

$$\begin{aligned} V^I(tr,tr) &= c + \delta^2 c + \delta^4 c + \cdots = \frac{c}{1-\delta^2}, \\ v^I(tr,tr) &= c. \end{aligned}$$

Now 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.

If tr is matched with $\tilde{d}$, then $D(tr, \tilde{d}) = D(\tilde{d}, tr) = 1$, and

$$\begin{aligned} V^I(tr, \tilde{d}) &= v^I(tr, \tilde{d}) = \ell, \\ V^I(\tilde{d}, tr) &= v^I(\tilde{d}, tr) = g. \end{aligned}$$

Thus, $\tilde{d}$ is a better response than tr if one expects to be matched with tr often.

Next we show the structure of the lifetime and average payoff of a player 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 in his lifetime as $V(s; p)$ and the average per period payoff s expects to receive during his lifetime as

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

A straightforward way to compute $V(s; p)$ is to set up a recursive equation. If p has a finite support, then we can write

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

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

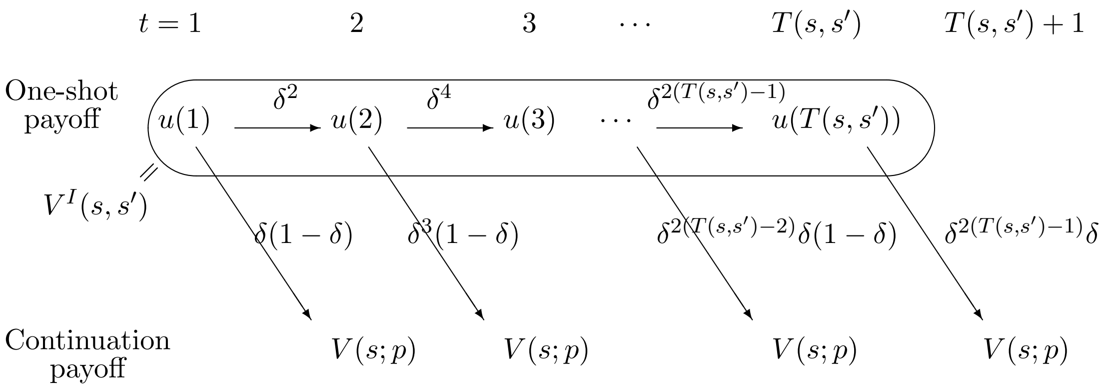

Figure 2: Transition structure of a payoff stream

By computation

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

Let

$$r(s;p) := \sum_{s' \in supp(p)} p(s')r(s,s').$$

Then

$$\begin{aligned} & V(s;p) = \sum_{s' \in supp(p)} p(s')V^I(s,s') + \{1 - r(s;p)\}V(s;p), \\ \iff \quad & V(s;p) = \sum_{s' \in supp(p)} \frac{p(s')}{r(s;p)} V^I(s,s'). \end{aligned}$$

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

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

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

where the ratio $r(s, s')/r(s; p)$ 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*[4]consisting of a single strategy s', then

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

Example 2. Consider a two-strategy distribution p such that α of the population use tr-strategy and $(1 - \alpha)$ of the population use $\tilde{d}$-strategy. The average payoff of tr-strategy when the strategy distribution in the matching pool is always p is:

$$\begin{aligned} v(tr; p) &= \frac{\alpha r(tr, tr) v^I(tr, tr) + (1-\alpha) r(tr, \tilde{d}) v^I(tr, \tilde{d})}{\alpha r(tr, tr) + (1-\alpha) r(tr, \tilde{d})} \\ &= \frac{\alpha c/(1+\delta) + (1-\alpha)(1-\delta)\ell}{\alpha/(1+\delta) + (1-\alpha)(1-\delta)} = \frac{\alpha c}{\alpha + (1-\alpha)(1-\delta^2)} + \frac{(1-\alpha)(1-\delta^2)\ell}{\alpha + (1-\alpha)(1-\delta^2)}. \end{aligned}$$

2.2 Nash Equilibrium

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

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

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

2.3 Neutral Stability

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

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.

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

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

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

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

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

where we have abused the notation and used s both as a strategy $s \in \boldsymbol{S}$ and a monomorphic strategy distribution consisting only of s, i.e., $s \in \mathcal{P}(\boldsymbol{S})$.

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.

Def. A (stationary) strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$ is a *Neutrally Stable Distribution* 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 will be 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 $\epsilon \in (0, \bar{\epsilon})$,

$$v(s; (1-\epsilon)p + \epsilon s') \geq v(s'; (1-\epsilon)p + \epsilon s').$$

It can be easily seen that any neutrally stable distribution is a Nash equilibrium.

We need some assumptions to justify this definition. Suppose that there is a continuous evolutionary process (such as replicator dynamics or best response dynamics) behind this definition and we discretize the process to model VRPD. Mutation occurs seldom enough so that, within the time span in which stationary strategy distribution is formed, only single mutation can occur. On the other hand, death-birth process takes place sufficiently quickly so that the strategy which is newly created by a mutation forms a stationary distribution within a single period. At the same time, strategy updates occurs sufficiently slowly so that incumbents' strategy distribution remains the same after a single period. While we do not insist that the above definition is the best among we can imagine, it is tractable and justifiable.

3 Voluntarily Repeated Cooperative Equilibria

In this section we consider strategies that voluntarily keep a partnership and play C repeatedly. There are many such strategies and only some of them constitute a Nash equilibrium or neutrally stable distribution.

3.1 Monomorphic Strategy Nash Equilibria

It is easy to see that the monomorphic strategy distribution $p_{tr} \in \mathcal{P}(\boldsymbol{S})$ consisting only of the trigger strategy in Example 1 is not a Nash equilibrium. The average payoff of tr-strategy and $\tilde{d}$-strategy under p_{tr} are $v(tr; p_{tr}) = v^I(tr, tr) = c$, and $v(\tilde{d}; p_{tr}) = v^I(\tilde{d}, tr) = g$. Therefore $\tilde{d}$-strategy earns strictly larger average payoff than the trigger strategy. By a similar logic, any strategy that starts with C in the first period of a match cannot constitute a Nash equilibrium.

Lemma 1. *Any pure strategy that starts with C in $t = 1$ cannot constitute a monomorphic Nash equilibrium.*

Proof. Let s be a strategy that plays C in the first period of a match and p be the monomorphic strategy distribution of s. Clearly, $\tilde{d}$ earns g as the **average** payoff under p, which is the maximum possible payoff any strategy can expect in any circumstances. I.e., $\tilde{d} \in BR(p)$ and $s \notin BR(p)$, which proves the assertion. □

On the other hand, $p_{\tilde{d}}$ consisting only of $\tilde{d}$-strategy is a Nash equilibrium. Against $\tilde{d}$, any strategy must play one-shot PD. Therefore, 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 $\tilde{d}$.

However, if the population consists only of $\tilde{d}$-strategy, some cooperative strategies can invade. Consider a class of strategies that play D in the first period (like $\tilde{d}$) and keep the partnership and play C in $t \geq 2$ if the partnership is alive. A strategy in this class earns the same average payoff as $\tilde{d}$-strategy when it is matched with $\tilde{d}$ but higher average payoff when it is matched with the same strategy (they get c at least in the second period). Therefore the population of $\tilde{d}$-strategy is invaded by a strategy in this class.

Some strategies in this class may not constitute a monomorphic strategy Nash equilibrium. Clearly, it is not an equilibrium to play C repeatedly after $t = 1$ and keep the partnership, regardless of history.

Consider the following class of strategies, which keeps the partnership if and only if both partners played C after $t = 1$.

Def. Let c_1-strategy be as follows.

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

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

c_1-strategy is a monomorphic strategy Nash equilibrium, if playing D and ending the partnership in some period $t \geq 2$ does not give a higher average payoff. This is not always true since the gain $g - c$ of playing D instead of C may be large as compared to the cost of restarting a partnership. In the following, we investigate the necessary and sufficient condition for the following generalization of c_1-strategy to be a Nash equilibrium.

Def. For any $T = 1, 2, 3, 4, \ldots$, let c_T-strategy be as follows.

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

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

We call the first T periods of c_T-strategy as the *trust-building periods* and the periods afterwards as the *cooperation periods*. Once the trust is built, c_T-strategy plays C disciplinarily in the sense that it ends the match if the partner did not choose C.

Let the monomorphic strategy distribution consisting only of c_T-strategy be denoted as p_T. Suppose that the stationary strategy distribution in the matching pool is p_T. The average payoff of c_T-strategy under p_T 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:

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

Since $v(c_T; p_T) = v^I(c_T, c_T)$, the average payoff is

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

Suppose that a player with c_T-strategy is at the beginning of t-th period of a match, where $t \in \{1, 2, \ldots\}$. The average payoff that this player expects to receive in the rest of his lifetime, denoted as $v(c_T; p_T, t)$, is called the *continuation average payoff of lifetime* and the average payoff that he expects to receive in the match, denoted as $v^I(c_T, c_T, t)$, is called the *continuation average payoff in the match*. The latter increases as t increases, since there is less time to build trust. Once the trust is built ($t \geq T+1$), the continuation average payoff in the match is constant and is c. (See Figure 3.)

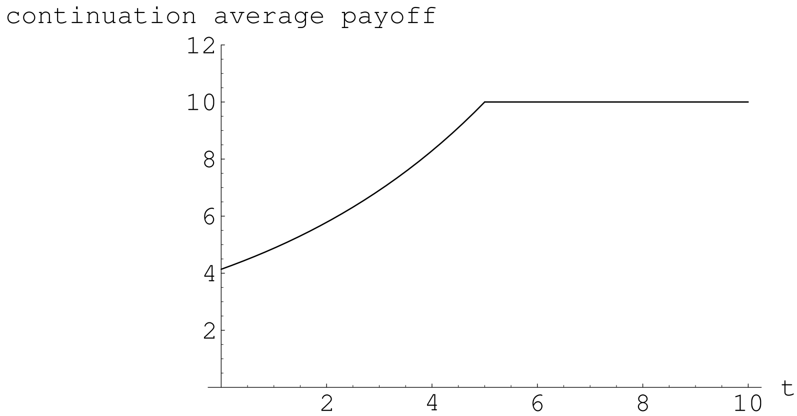

Figure 3: The continuation average payoff of c_T-strategy in the match with c_T.

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

The increasing nature of the continuation average payoff of c_T works as an incentive device to play C in the cooperation periods. A deviation from c_T in the cooperation periods costs the player *trust capital*, which is defined as the value difference between the continuation average payoff in $t > T$ and the average payoff in the matching pool. It is computed as follows.

By the definition,

$$v(c_T; p_T, t) = r(c_T, c_T, t)v^I(c_T, c_T, t) + [1 - r(c_T, c_T, t)]v(c_T; p_T) \tag{5}$$

where,

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

It follows then:

$$\begin{aligned} v(c_T; p_T, t) &= v(c_T; p_T) + r(c_T, c_T, t)[v^I(c_T, c_T, t) - v(c_T; p_T)] \\ &= v(c_T; p_T) + r(c_T, c_T, t) \cdot TC(c_T, c_T, t) \end{aligned} \tag{6}$$

where:

$$TC(c_T, c_T, t) = \begin{cases} \delta^{2(T-t)}(1-\delta^{2t})(c-d) & \text{if} \quad t \leq T \\ (1-\delta^{2T})(c-d) & \text{if} \quad t \geq T+1 \end{cases} \tag{7}$$

which is the trust capital.

For any deviation strategy s which imitates c_T until some period in the cooperation periods and plays D, the one-shot payoff is at most g but the player must return to the matching pool next period and settle with the average payoff of $v(c_T; p_T)$ during the periods when partnership would have been maintained (for the portion of $r(c_T, c_T, t)$ of his lifetime). The resulting average payoff is:

$$\begin{aligned} v(s, c_T; t) &= r(c_T, c_T, t)[(1-\delta^2)g + \delta^2 v(c_T; p_T)] + [1 - r(c_T, c_T, t)]v(c_T; p_T) \\ &= v(c_T; p_T) + r(c_T, c_T, t)(1-\delta^2)[g - v(c_T; p_T)]. \end{aligned} \tag{8}$$

It follows that the *incentive constraint* for c_T-strategy to be followed during the cooperation periods is (6) being not less than (8), *i.e.,*

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

Or,

$$\delta^2 \frac{1-\delta^{2T}}{1-\delta^2}(c-d) \geq g - c. \tag{9}$$

It follows immediately that, as long as $g - c > 0$, a positive number of trust building periods is necessary for equilibrium cooperation (i.e., $T \geq 1$) as was already shown in Lemma 1.[5]

Now we prove that in fact the above incentive constraint is the only condition that is required for p_T to be a Nash equilibrium. Let *on-path history* at a decision node of $t = 2, 3, \ldots$, 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 2. *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.*

[5] By contrast, if there is a matching friction as in Shapiro-Stiglitz (1984) so that one might have to spend some periods without forming a partnership, it is possible to play C from the beginning for sufficiently low probability of getting a new match. For the details, see Section 4.

(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 in some period $t \leq T$ 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

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

Proof. See Appendix B.

The sufficient length of trust-building periods T depends on the payoff configuration $G = (g, c, d, \ell)$ and the death probability δ. For each $(\delta, T) \in (0, 1) \times \{1, 2, \ldots\}$, define

$$f(\delta, T) := \delta^2 \frac{1 - \delta^{2T}}{1 - \delta^2}.$$

For each G and $T = 1, 2, \ldots$, define $\underline{\delta}_G(T)$ implicitly by

$$f(\underline{\delta}_G(T), T) = \frac{g - c}{c - d}.$$

Then the incentive constraint (9) is satisfied for (δ, T, G) if and only if

$$\delta \geq \underline{\delta}_G(T).$$

Because $f(\delta, T)$ is increasing in T, $\underline{\delta}_G(T)$ is decreasing in T. It is easy to see that $f(\delta, 1) = \delta^2$ and $\lim_{T \to \infty} f(\delta, T) = \delta^2/(1 - \delta^2)$. Therefore, for any G,

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

Since $\underline{\delta}_G(\infty) < 1$,for any G and for any $\delta \in (\underline{\delta}_G(\infty), 1)$, we can define the minimum length $\underline{T}_G(\delta)$ of trust building periods that satisfies the incentive constraint by

$$\underline{T}_G(\delta) := argmin_{T \in \{1,2,\ldots\}} \{\underline{\delta}_G(T) \mid \delta \geq \underline{\delta}_G(T)\}.$$

(See Figure 5 in Section 3.3.)

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 $T \geq \underline{T}_G(\delta)$.*

Proof. Lemma 2 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 (9) holds, i.e., $T \geq \underline{T}_G(\delta)$. Strategies that differ from c_T-strategy off the play path do not give a higher payoff. □

Proposition 1 shows that for sufficiently long trust-building periods and sufficiently high survival probability, voluntarily repeated cooperation is sustained.

Note that the lower bound to the discount factor that sustains the trigger-strategy equilibrium if there is no option to end a match is $\sqrt{\frac{g-c}{g-d}} = \underline{\delta}_G(\infty)$. This means that cooperation in voluntarily repeated PD requires more patience.

3.2 Monomorphic NSS

As we discussed at the beginning of Section 3.1., the myopic strategy that plays D and ends the partnership immediately is not NSS.

Lemma 3. *Myopic $\tilde{d}$-strategy played by all players is not an NSS.*

Proof.: For any $\epsilon \in (0,1)$, let $p := (1-\epsilon)p_{\tilde{d}} + \epsilon c_1$. From (1),

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

Since $v^I(c_1,\tilde{d}) = d$ and $v^I(c_1,c_1) = (1-\delta^2)d + \delta^2 c > d$, c_1-strategy invades $p_{\tilde{d}}$. □

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. See Appendix B.

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.[6] The other kind is the strategies that play D in the

[6] However, a strategy distribution may drift toward such strategies and eventually force the distribution out of NSS. Since this concern is of dynamic nature, we postpone the discussion til Section 5.

cooperation periods. When $\delta > \underline{\delta}_G(T)$, however, Lemma 2 (c) implies that such strategies are not alternative best reply. Therefore c_T-strategy is NSS for this case.

There remains the boundary case of $\delta = \underline{\delta}_G(T)$ that p_T is a Nash equilibrium. For this case, we consider an alternative best reply to p_T which earns the highest payoff when meeting itself. Since we only allow pure strategy entrants, without loss of generality we can focus on c_{T+1}-strategy as the alternative best reply that earns the highest payoff against itself. Below we identify a sufficient condition that c_{T+1}-strategy cannot invade p_T.

Let $p_T^{T+1}(\alpha)$ be the two-strategy distribution consisting of c_T and c_{T+1}-strategies such that α of the players are c_T-strategy in the matching pool. Note that c_{T+1}-strategy cannot invade p_T if and only if there exists $\bar{\alpha} \in (0,1)$ such that for any $\alpha \in (\bar{\alpha}, 1]$, $v(c_T; p_T^{T+1}(\alpha)) \geq v(c_{T+1}; p_T^{T+1}(\alpha))$.

For any G, any $\alpha \in [0,1]$ and any $\delta \in (0,1)$, let

$$\Delta v_T(\alpha, \delta) = v(c_T; p_T^{T+1}(\alpha)) - v(c_{T+1}; p_T^{T+1}(\alpha)).$$

The following "characteristic" function to Δv_T makes the computation simple.

Lemma 5. *For any G, any $\alpha \in [0,1]$ and any $\delta \in (0,1)$, let*

$$\Delta \tilde{v_T}(\alpha, \delta) = \{v(c_T; p_T^{T+1}(\alpha)) - v(c_{T+1}; p_T^{T+1}(\alpha))\}\{1 - (1-\alpha)\delta^{2(T+1)}\}(1 - \alpha\delta^{2(T+1)})/\delta^{2T}.$$

Then $\Delta v_T(\alpha, \delta) \geq 0$ if and only if $\Delta \tilde{v_T}(\alpha, \delta) \geq 0$. Moreover,

(a) For each $\delta \in (0,1)$, $\Delta \tilde{v_T}(\alpha, \delta)$ is a concave, quadratic function of α; and

(b) For any $\delta \in (0,1)$, $\Delta \tilde{v_T}(0, \delta) < 0$.

Proof.: This is by computation. See Appendix B.

When $\alpha = 1$, by definition of $\underline{\delta}_G(T)$, $\delta > \underline{\delta}_G(T)$ if and only if $\Delta \tilde{v_T}(1, \delta) > 0$. Therefore for any $\delta > \underline{\delta}_G(T)$, c_{T+1} cannot invade p_T.

By definition of $\underline{\delta}_G(T)$, $\Delta v_T(1, \underline{\delta}_G(T)) = \Delta \tilde{v_T}(1, \underline{\delta}_G(T)) = 0$. We argue indirectly by the concavity of $\Delta \tilde{v_T}$ function and Lemma 5 (b) that c_{T+1}-strategy cannot invade p_T if and only if $\Delta \tilde{v_T}(\alpha, \delta)$ has a strictly negative slope at $\alpha = 1$. This implies that for α near 1, $\Delta \tilde{v_T}(\alpha, \delta) > 0$, which is equivalent to $\Delta v_T(\alpha, \delta) > 0$. See Figure 4.

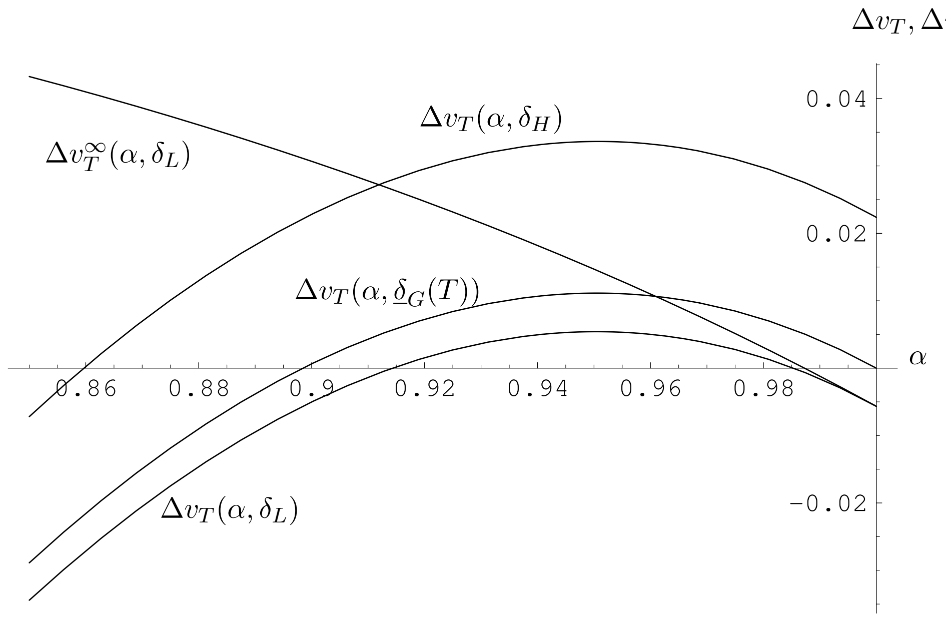

Figure 4: The value difference between c_T-strategy and c_{T+1}-strategy under two-strategy distribution p_T^{T+1} and geometric strategy distribution $p_T^\infty(\alpha)$ (see Section 3.3) (Parameter values: $g = 10, c = 6, d = 1, \ell = -1, T = 1, \delta_L = 0.893 < \underline{\delta}_G(T) \approx 0.894427 < \delta_H = 0.9$.)

Define $\hat{\delta}_G(T)$ implicitly as follows:

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

Lemma 6. *For any G and any T, $\frac{\partial \Delta \tilde{v_T}}{\partial \alpha}(1, \underline{\delta}_G(T)) < 0$ if and only if*

$$\delta > \hat{\delta}_G(T). \tag{10}$$

Proof. See Appendix B.

Since $r(c_T, c_T) = (1-\delta)\{1 + \delta^2 + \cdots\}$ and $r(c_{T+1}, c_T) = (1-\delta)\{1 + \delta^2 + \cdots + \delta^{2T}\}$, the condition (10) is equivalent to

$$(c - d)r(c_T, c_T) > (g - \ell)r(c_{T+1}, c_T).$$

The LHS can be interpreted as the relative merit of c_T-strategy against c_{T+1}-strategy (to start co-operating one period early when meeting itself) and the RHS is the relative merit of c_{T+1}-strategy (when meeting the other strategy). As T increases, (10) becomes more difficult to satisfy, since the difference between $r(c_T, c_T)$ and $r(c_{T+1}, c_T)$ becomes smaller. Hence (10) puts an upper bound to T. Figure 4 shows, however, the existence of G and $T \geq T_G^*$ that satisfy (10).

In sum, we have the following parametric condition for the existence of monomorphic NSS.

Proposition 2. *(a) For any G, any $\delta > \underline{\delta}_G(\infty)$ such that $\delta \neq \underline{\delta}_G(T)$ for any T, p_T is neutrally stable for any $T \geq \underline{T}_G(\delta)$.*

(b) For any G, any $\delta > \underline{\delta}_G(\infty)$ such that $\delta = \underline{\delta}_G(T)$ for some T, p_T is neutrally stable if and only if $T \geq \underline{T}_G(\delta)$ and $\delta > \hat{\delta}_G(T)$.

We have shown that when δ is large enough, voluntarily repeated cooperation is NSS with sufficient number of trust-building periods, even though myopic defection is not.

3.3 Polymorphic Strategy Distributions

In this section we consider the possibility that different strategies co-exist in an equilibrium. There are two types of diversity. We first consider c_T-strategies with different length of trust building. We then consider strategy combination in which (C, D) and (D, C) are played after trust building.

3.3.1 Different Trust-Building Periods

We show the existence of polymorphic neutrally stable distributions, in which various c_T-strategies co-exist. Diversity of trust-building periods generates subtle balances so that each strategy benefits from matches with certain strategies while the same strategy loses by matcthes with other strategies. In an equilibrium, if population share of a certain strategy changes, payoff balance is destroyed, forcing a chain of population adjustments to take place which, in turn, brings the strategy distributuion back to the original equilibrium.

If a polymorphic strategy distribution with the infinite support is a candidate of a neutrally stable distribution, then the population distribution of c_t-strategies must be "geometric".

Lemma 7. *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 B.

Moreover, if $\{c_T, c_{T+1}, \ldots\}$ are distributed according to the geometric distribution and c_T and c_{T+1} have the same average payoff, then all other strategies in the support have also the same payoff.

Lemma 8. *Take any G and $T < \infty$. Let $p_T^\infty(\alpha)$ be a stationary strategy distribution with the support $\{c_T, c_{T+1}, \ldots\}$ such that the fraction of $c_{T+\tau}$-strategy is $\alpha(1-\alpha)^\tau$ for each $\tau = 0, 1, 2, \ldots$. Then*

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

Proof: See Appendix B.

By a similar logic to Lemma 2, strategies which end the partnership earlier than T (either by choosing e or by choosing C along the on-path history) cannot invade $p_T^\infty(\alpha)$, nor strategies which end the partnership after playing C in the cooperation periods. It remains to find out the condition on T and δ that warrants the existence of $\alpha^* \in (0, 1)$ which satisfies the following three conditions.

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

   This condition, together with Lemma 8, implies that all strategies in the support of $p_T^\infty(\alpha^*)$ have the same payoff.

2. for $\alpha < \alpha^*$, $v(c_T; p_T^\infty(\alpha)) > v(c_{T+1}; p_T^\infty(\alpha))$ and $\alpha > \alpha^*$, $v(c_T; p_T^\infty(\alpha)) < v(c_{T+1}; p_T^\infty(\alpha))$ at least near α^*. (See Figure 4.)

   This implies that c_T or c_{T+1} cannot invade (increase the fraction in) $p_T^\infty(\alpha^*)$. Moreover, if $c_{T+\tau}$ ($\tau \geq 2$) increases the fraction in $p_T^\infty(\alpha^*)$, it is as if c_{T+1} increases. Hence c_T gets larger payoff and thus $c_{T+\tau}$ ($\tau \geq 2$) cannot invade in the sense of increasing the fraction.

3. Incentive constraints to prevent D once the cooperation periods started for each of $c_{T+\tau}$-strategy ($\tau = 0, 1, 2, \ldots$): for each τ, 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

$$v(c_{T+\tau}; p_T^\infty(\alpha^*)) > v(s_\tau; p_T^\infty(\alpha^*)). \tag{11}$$

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

Lemma 9. *Take any G such that there exists $T \in \{1, 2, \ldots\}$ such that $\hat{\delta}_G(T) < \underline{\delta}_G(T)$. For any T 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 $\delta \geq \delta^*_G(T)$, there exists $\alpha^*(\delta)$ that satisfies 1 and 2.*

Proof: See Appendix B.

To see if such G exists, note that for any G, $\hat{\delta}_G(T)$ is increasing in T. $\hat{\delta}_G(0) = \sqrt{1 - \frac{c-d}{g-\ell}} < \infty = \underline{\delta}_G(0)$, 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. (See Figure 5.)

Lemma 10. *For any (G, T) and for δ sufficiently close to but less than $\underline{\delta}_G(T)$,*

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

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

Proof: See Appendix B.

By combining Lemmas 9 and 10, we have the existence of a polymorphic NSD with the support $\{c_T, c_{T+1}, \ldots\}$.

Proposition 3. *Take any G such that there exists $T \in \{1, 2, \ldots\}$ such that $\hat{\delta}_G(T) < \underline{\delta}_G(T)$. For any T such that $\hat{\delta}_G(T) < \underline{\delta}_G(T)$, for $\delta \in (\delta^*_G(T), \underline{\delta}_G(T))$ that is sufficiently close to $\underline{\delta}_G(T)$, there is a neutrally stable polymorphic strategy distribution of the form $p_T^\infty(\alpha^*(\delta))$ for some $\alpha^*(\delta) \in (0, 1)$.*

Proof: Lemma 9 shows the existence of $\alpha^*(\delta)$ such that all strategies in the support have the same payoff and that no strategy in the support can increase its fractions. Lemma 10 shows that strategies that differ on the play path during the cooperation periods of any $c_{T+\tau}$ strategy cannot earn higher payoff. To finish, we show that other strategies that differ on the play path from $\{c_T, c_{T+1}, \ldots\}$ do not earn higher payoff, in Appendix B. □

Therefore, cooperation and exploitation can co-exist. Moreover, it is possible that the polymorphic NSD includes c_T with shorter trust-building periods than any monomorphic NSS under the relevant

(G, δ). The idea is as follows. Let $T+1$ be the shortest trust-bulding periods of monomorphic NSS. Since c_T itself is not NSS, another strategy that chooses D during the cooperation periods can enter the population. Among mutants, c_{T+1} strategy that chooses D at $T+1$ earns the highest payoff when meeting itself. However, if c_{T+1} met another c_{T+1}, there is an incentive to play D at $T+2$. And so on. Therefore $c_{T+1}, c_{T+2}, \ldots$ will enter. If the distribution becomes stable, it must be the case that (a) the distribution must be geometric (Lemma 7) and (b) as the share α of the most efficient strategy c_T decreases, the exploiting strategy c_{T+1} should have less merit. The property (b) is represented by the condition $\hat{\delta}_G(T) < \delta < \underline{\delta}_G(T)$ in Lemma 9 that the value difference between c_T and c_{T+1} is decreasing and negative near $\alpha = 1$.

Figure 5 below illustrates the parameteric conditions of monomorphic NSS and polymorphic NSD.

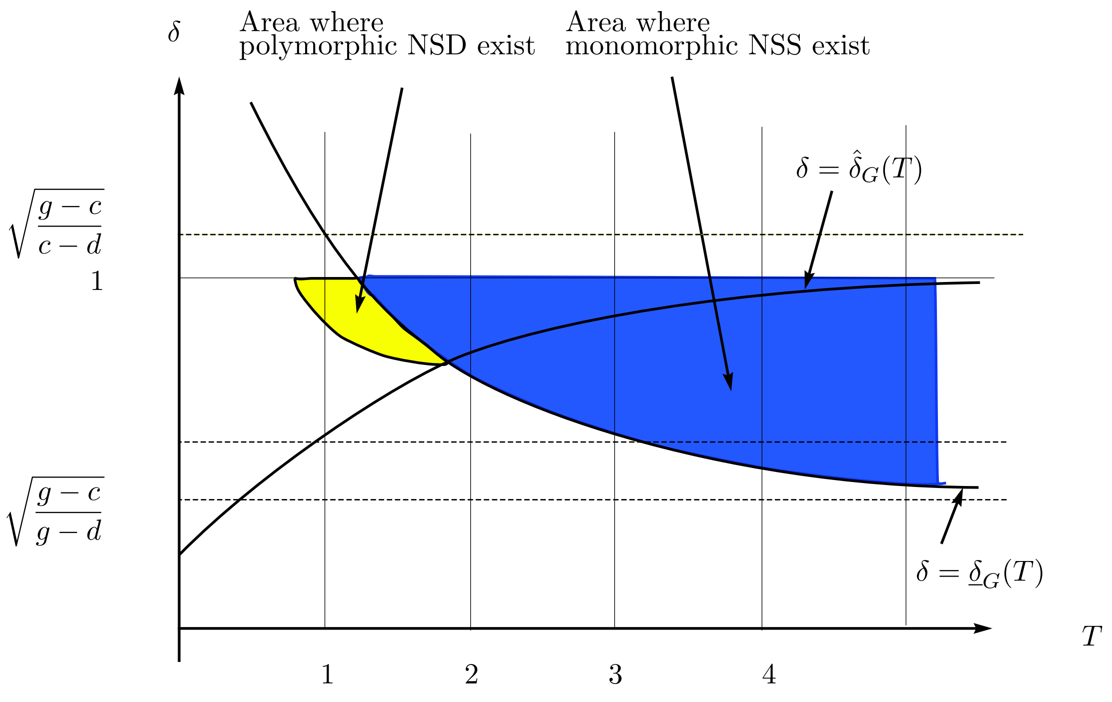

Figure 5

3.3.2 Off-diagonal Coordination

We consider strategy distributions in which (C, D) and (D, C) are played after trust building. For some G, the average payoff of repeating (C, D) and (D, C) is higher than the repetition of (C, C). Therefore it is important to consider this class.

To be filled later.

3.4 Welfare Comparison

Since there are multiple cooperative equilibria, one may want to find out the most efficient one. Among the same type (monomorphic or geometric polymorphic) NSD, clearly the one with the shortest trust-building periods is most efficient (but not fully efficient since $T > 0$.) However, it is difficult to compare a monomorphic NSS with a polymorphic NSD. With the same trust-building periods T, p_T gives higher average payoff than each strategy in the support of $p_T^{\infty}(\alpha^*)$, since c_T in the latter is exploited by other strategies in the support. However, there may be a polymorphic NSD with a shorter trust-building periods than the most efficient monomorphic NSS as we discussed above.

4 Matching Friction and Efficiency Wage

Trust capital works as an incentive constraint because, if one deviates from c_T by playing D in a cooperation period, s/he is forced to start a new partnership wasting next T periods for trust building before cooperation begins. This loss in payoff deters a deviation incentive.

Note, however, this logic holds because we implicitly assumed that one can always find a new partnership in the matching pool. In fact, if one assumes that one can find a new parnership only with the probability $1 - u \in (0, 1)$ and spends the next period without a partner and payoff of 0 with probability $u \in (0, 1)$, incentive constraint may be satisfied without trustbuilding periods.

With this possibility of "unemployment", average payoff that c_T strategy player expects to receive in the matching pool (but before s/he finds a partner) is:

$$\begin{aligned} v^0(c_T; p_T) &= (1-u)v(c_T; p_T) \\ &= (1-u)[(1-\delta^{2T})d + \delta^{2T}c] \\ &= v(c_T; p_T) - U_T, \end{aligned}$$

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), and $U_T = u[(1-\delta^{2T})d + \delta^{2T}c]$.

It follows that the value of c_T in the t-th period of partnership is now defined as:

$$\begin{aligned} v(c_T; p_T, t) &= r(c_T, c_T, t)v(c_T, c_T, t) + [1 - r(c_T, c_T, t)]v^0(c_T; p_T) \\ &= v^0(c_T, p_T) + r(c_T, c_T, t)[v(c_T, c_T, t) - v^0(c_T; p_T)] \\ &= v^0(c_T; p_T) + r(c_T, c_T, t)TC^0(c_T, c_T, t) \end{aligned}$$

where,

$$TC^0(c_T, c_T, t) = \begin{cases} \delta^{2(T-t)}(1-\delta^{2t})(c-d) + u[(1-\delta^{2T})d + \delta^{2T}c] & \text{if} \quad t < T \\ (1-\delta^{2T})(c-d) + u[(1-\delta^{2T})d + \delta^{2T}c] & \text{if} \quad t \geq T \end{cases}$$

or,

$$\begin{aligned} TC^0(c_T, c_T, t) &= TC(c_T, c_T, t) + u[(1-\delta^{2T})d + \delta^{2T}c] \\ &= TC(c_T, c_T, t) + U_T. \end{aligned}$$

Hence, if $d \geq 0$, trust capital is larger with unemployment than without it. In fact, if:

$$\begin{aligned} TC^0(c_T, c_T, T) &= TC(c_T, c_T, t) + U_T \\ &\geq (1-\delta^2)[g - v^0(c_T; p_T)] \\ &= (1-\delta^2)[g - v(c_T; p_T) + U_T], \end{aligned}$$

or equivalently:

$$(1-\delta^{2T})\delta^2(c-d) \geq (1-\delta^2)(g-c) - \delta^2 U_T,$$

then IC constraint is satisfied.

In fact, if:

$$\begin{aligned} \delta^2 U_T &= u[(1-\delta^{2T})d + \delta^{2T}c] \\ &\geq (1-\delta^2)(g-c), \end{aligned}$$

or, equivalently,

$$u \geq \frac{(1-\delta^2)(g-c)}{(1-\delta^{2T})d + \delta^{2T}c},$$

then IC constraint is satisfied even for c_0 and cooperation without trust building period becomes a self-sustaining state.

As readers may have observed already, this is the well known result of efficiency wage theory (see, e.g., Shapiro and Stiglitz, 1984, and Okuno-Fujiwara, 1989) where unemployment works as a disciplinary device that deters moral hazard behavior. Trust building periods in our model works as an alternative disciplinary device against moral hazard.

5 Equivalent Strategies, Drift, and Equilibrium Evolutionary Stability

5.1 Equivalent Strategies and Drift

Our discussions so far implicitly assumed that there is a single pair of strategies corresponding to each matching game outcome. However, because our matching game itself is an extensive form game, there are many (pure) strategies that generate the same outcome. Thus, each Nash equilibrium outcome corresponds to a set of strategy distributions, not to a single strategy distribution.

To make it precise, we define equivalent strategies as follows.

Def. Given a strategy distribution, $p \in \mathcal{P}(\boldsymbol{S})$, a strategy $s' \in \boldsymbol{S}$ is *equivalent with strategy* $s \in \text{supp}(p)$ *in distribution* p if, for all $s" \in \text{supp}(p)$, the matching game outcome between s' and $s"$ is identical with that between s and $s"$.

We write $s' \approx_p s$ when s' is equivalent with s in p. Given $s \in \text{supp}(p)$, we shall denote the set of all (pure) strategies which are equivalent with s in p as $E(s; p)$, or:

$$E(s; p) := \{s' \in \boldsymbol{S} | s' \approx_p s\}.$$

We shall denote the set of all (pure) strategies which are equivalent with some $s \in \text{supp}(p)$ in p as $E(p)$, or:

$$E(p) := \{s' \in \boldsymbol{S} | \exists s \in \text{supp}(p) \text{ such that } s \approx_p s'\}.$$

Then, given a strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$, we can define the set of strategy distributions which are equivalent with p as follows:

$$\mathcal{P}(E(p)) := \{p' \in \mathcal{P}(\boldsymbol{S}) | \text{supp}(p') \subseteq E(p) \text{ and outcome distributions of } p \text{ and } p' \text{ are identical}\}.$$

Def. For each Nash equilibrium that consists of a single strategy distribution, $p \in \mathcal{P}(\boldsymbol{S})$, an associated *set of NE strategy distributions* (*NE set or NE component*) is:

$$NE(p) := \{p' \in \mathcal{P}(E(p)) | BR(p') \subseteq E(p)\},$$

Remark 1. *If* $p \in \mathcal{P}(\boldsymbol{S})$ *is a monomorphic equilibrium NE,* $NE(p)$ *is a monomorphic outcome NE. If* $p \in \mathcal{P}(\boldsymbol{S})$ *consists of non-equivalent strategies in* p*,* $NE(p)$ *is a polymorphic outcome NE.*

We shall denote by $NE(\mathcal{P}(\boldsymbol{S}))$ the set of NE Sets of the VRPD.

We define similar concept for neutral stability. However, we first generalize the concept of invasion as follows.

Def. Given $\epsilon > 0$ and a distribution $p \in \mathcal{P}(\boldsymbol{S})$, a strategy distribution $p' \in \mathcal{P}(\boldsymbol{S})$ *invades* p if for any $s' \in \text{supp}(p')$ and $s \in \text{supp}(p)$,

$$v(s'; (1-\epsilon)p + \epsilon p') \geq v(s; (1-\epsilon) + \epsilon p'),$$

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

$$v(s'; (1-\epsilon)p + \epsilon p') > v(s; (1-\epsilon) + \epsilon p').$$

Def. A set of Nash Equilibrium strategy distributions, $NE(p)$, is a *Neutrally Stable Set* if, for any $p \in NE(p)$ and for any $p' \in \mathcal{P}(\boldsymbol{S})$, there exists $\varepsilon \in \bar{(}0,1)$ such that p' cannot invade p for any $\varepsilon \in (0, \bar{\varepsilon})$.

A neutrally stable set (NSSet) is called a *monomorphic outcome NSSet* if its outcome is monomorphic while it is called a *polymorphic outcome NSSet* if its outcome is polymorphic.

5.2 Example and Drift

An illustrative example may be useful. Let us restrict the set of strategies to $\mathcal{P}(\boldsymbol{S}) = \{c_1, \hat{c}_1, c_\infty, \tilde{d}\}$. We have already defined c_1, which is a disciplinary strategy with trustbuilding period of length 1, and $\tilde{d}$, which is a myopic strategy to play D once and then end the partnership.

By definition, c_∞ is the strategy defined as:

For any $t = 1, 2, \ldots$: play D and k for any observation.

Finally, $\hat{c}_1$ is defined as follows:

$t = 1$: Play D and k for any observation.

$t \geq 2$: Play C and k for any observation.

Now consider the monomorphic strategy distribution p_1 which consists only of the strategy c_1. It is clear that $\hat{c}_1 \approx_{p_1} c_1$ and $E(p_1) = \{c_1, \hat{c}_1\}$ as $\hat{c}_1$'s behavior with c_1 and with $\hat{c}_1$ is identical to the behavior of c_1 who matches with another c_1. It follows that

$$v(c_1; p) = v(\hat{c}_1; p) = (1 - \delta^2)d + \delta^2 c \text{ for all } p \in \mathcal{P}(E(p_1)).$$

However, notice:

$$\begin{aligned}
v^I(c_\infty, c_1) &= d, & r(c_\infty, c_1) &= \frac{(1-\delta^4)/(1-\delta^2)}{1/(1-\delta)} = \frac{1-\delta^4}{1+\delta}, \\
v^I(c_\infty, \hat{c}_1) &= (1-\delta^2)d + \delta^2 g, & r(c_\infty, \hat{c}_1) &= \frac{1/(1-\delta^2)}{1/(1-\delta)} = \frac{1}{1+\delta}, \\
v^I(\hat{c}_1, c_\infty) &= (1-\delta^2)d + \delta^2 \ell, & r(\hat{c}_1, c_\infty) &= \frac{1/(1-\delta^2)}{1/(1-\delta)} = \frac{1}{1+\delta}.
\end{aligned}$$

In order to find $NE(p_1)$, let $p \in \mathcal{P}(E(p_1))$ such that $p(\hat{c}_1) = 1 - \alpha$ while $p(c_1) = \alpha$.

$$\begin{aligned}
v(c_1; p) = v(\hat{c}_1; p) &= (1-\delta^2)d + \delta^2 c, \quad r(c_1, c_1) = r(\hat{c}_1, c_1) = r(\hat{c}_1, \hat{c}_1) = \tfrac{1}{1+\delta}, \\
v(c_\infty; p) = (1-\mu(\alpha))v^I(c_\infty, c_1) + \mu(\alpha)v^I(c_\infty, \hat{c}_1), & \qquad \mu(\alpha) = \tfrac{\alpha r(c_\infty, \hat{c}_1)}{\alpha r(c_\infty, \hat{c}_1) + (1-\alpha) r(c_\infty, c_1)}.
\end{aligned}$$

Then straightforward computations yield:

$$v(c_1; p) \geq v(c_\infty; p) \Leftrightarrow \alpha \leq \bar{\alpha} := \frac{(1-\delta^4)(c-d)}{g - c + (1-\delta^4)(c-d)}.$$

It follows then:

$$NE(p_1) = \{p \in \mathcal{P}(E(p_1)) | p(\hat{c}_1) \leq \bar{\alpha}\} \text{ while } NS(p_1) = \{p \in \mathcal{P}(E(p_1)) | p(\hat{c}_1) < \bar{\alpha}\}.$$

Thus, we find the following dynamics in this four-strategies population.

One may think of the starting strategy distribution which consists only of $\tilde{d}$, or $p_{\tilde{d}} \in \mathcal{P}(\boldsymbol{S})$. Clearly, with the current definition of $\mathcal{P}(\boldsymbol{S})$, $p_{\tilde{d}}$ is the only component of $ES(p_{\tilde{d}})$ or $\{p_{\tilde{d}}\} = ES(p_{\tilde{d}})$. As we have shown, however, $ES(p_{\tilde{d}})$ is not a NSSet and c_1 (possibly along with some fraction of $\hat{c}_1$) invades. if, moreover, IC condition is satisfied, then c_1 or its equivalent strategies start to propagate. This process eventually results to a distribution $p \in NS(p_1)$, which is a NSSet.

However, strategy distribution may drift as c_1 and $\hat{c}_1$ yield identical expected payoffs as long as strategy distribution stays within $\mathcal{P}(E(p_1))$. Sooner or later, this drift will move strategy distribution out of $NS(p_1)$. If it occurs, c_∞ invades and starts to propagate in the society. A monomorphic distribution, p_∞, consisting only of c_∞ may evolve.

However, in p_∞, strategies c_∞ and $\tilde{d}$ are equivalent and population drift may occur, making strategy distribution with sufficiently many $\tilde{d}$, to which c_1 may invade. Thus the soecity's strategy distribution perpetually cycles along these states without resting forever.

5.3 Equilibrium Evolutionarily Stable Set

Neutrally stable strategies (or NS distributions) have been criticized on two grounds.

First, it is too weak as a solution concept because it allows drift. As our example in the previous subsection shows, given a candidate strategy for equilibrium, there may be many other strategies that generate the same payoffs because their behaviors are different from candidate strategy only at off-path nodes. If such strategies invade, eventually, stability of candidate strategy (distribution) may be destroyed.

Second, in a human society, no new strategies will propagate even if they are better response to the incument strategies. Humans are more intelligent and they should adopt such new strategies if they are not only better response to the incumbent strategies but also they remain best response after new strategies propagate. Simply put, invading strategies should be restricted to those strategies which are best responses against the "post-entry" distributions.

These considerations lead to the concept of "Equibrium Evolutionarily Stability", which is originally introduced by Swinkles (1992). Formally,

Def. A set $\Theta \sqsubseteq \mathcal{P}(\boldsymbol{S})$ is *equilibrium evolutionarily stable (EES)* if it is minimal with respect to:

1. Θ is closed,
2. $\Theta \subseteq NE(\mathcal{P}(\boldsymbol{S}))$,
3. there exists $\bar{\varepsilon} \in (0,1)$ such that for all $\varepsilon \in (o, \bar{\varepsilon})$ such that, for all $p \in \Theta$ and for all $p' \in \mathcal{P}(\boldsymbol{S})$:

$$\mathrm{supp}(p') \subseteq BR((1-\varepsilon) + \varepsilon p') \Rightarrow (1-\varepsilon)p + \varepsilon p' \in \Theta.$$

6 Refinement by Cheap Talk

Recall that the reason that c_T-strategy can invade the population of $\tilde{d}$-strategy is that c_1-strategy keeps the partnership after (D, D) as a signal for future cooperation. 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.

We now introduce cheap talk to our model.

6.1 Model

Assume that when two players are newly matched, they simultaneously choose a message from a countably infinite set M. M is common to all players. The messages do not alter the payoff and thus are cheap-talk. We assume that the message choices of the matched players is private information in the match.

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

1. $m \in M$ specifies an initial message,
2. $\sigma : M \to \boldsymbol{S}$ specifies a strategy $\sigma(m')$ chosen for each message $m' \in M$ the partner announces.

In the rest of the paper, however, in order to ease the notation we shall denote pure strategy by σ, omitting the message m. 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".

Let $\boldsymbol{S}^{CT}$ be the set of all pure strategies of VRPD with cheap talk. There are many ways to extend a pure strategy in $\boldsymbol{S}$ (of the ordinary VRPD) into the cheap-talk model.

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 result of information exchange. For example, 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.

6.2 Evolutionary Stability of Babbling Equilibria

When we discuss babbling strategies, for each pure strategy $s \in \boldsymbol{S}$, we write corresponding babbling strategy of the cheap talk model (actually, set of strategy distributions because message choice is arbitrary) as $\sigma^B(s) \in \mathcal{P}(\boldsymbol{S}^{CT})$. Because babbling strategies ignores messages from partners and the message choice is irrelevant for the entire population, we can also 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 in an obvious manner. In order to ease notations, we shall write this distribution (actually, set of distributions because message choice is arbitrary) as $\sigma^B(p) \in \mathcal{P}(\boldsymbol{S}^{CT})$, with the superscript B.

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

As is well-known in this research area, babbling extension of Nash equilibria are Nash equilibria of the cheap-talk model.

Lemma 11. *For any Nash equilibrium strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$ of VRPD, the associated babbling strategy distribution $p^B \in \mathcal{P}(\boldsymbol{S}^{CT})$ is a Nash equilibrium strategy distribution of the cheap talk model.*

In subsection 3.4.entitled "Welfare Comparison", we showed that, given G, either the monomorphic NSS with the minimum trust-building periods or one of the polymorphic 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 $p^{*B} \in \mathcal{P}(S^{CT})$ be the associated babbling strategy. Clearly, there is an associated set of NE strategy distribution, $NE(p^{*B}) \in NE(\mathcal{P}(\boldsymbol{S}))$.

Lemma 12. *For any G, the NE outcome distributions $NE(p^{*B}) \in NE(\mathcal{P}(\boldsymbol{S}))$ is the Equlibrium Evolutionarily Stable Set.*

Proof. See Appendix B.

7 Concluding Remarks

7.1 Related Literature

Datta (1996) and Kranton (1996a) consider a complete information, two-player, voluntarily repeated game similar to ours. The component game is a continuum-action prisoner's dilemma, representing a borrower-lender situation etc. Therefore the players in their model can gradually increase the "level of cooperation", which makes the same disciplining system as our c_T-strategy.

Ghosh and Ray (1996) consider an incomplete information model. Since there are players who always defect, it is rational to play D at the beginning of a partnership, not only to build trust but also to distinguish the types of the opponents. Their equilibrium notion requires that the joint deviation by the partners should not improve the payoffs (in addition to the individual deviations) and thus if such equilibrium exists, it is efficient. However, the existence is not warranted.

These models assume full rationality of players and adopt equilibrium notions requiring sequential rationality.

Carmichael and MacLeod (1997) formulate an evolutionary model with initial gift exchange stage added to the voluntarily repeated prisoner's dilemma. The gift choice works the same way as the "level of cooperation" adjustment.

Overall, the idea of trust-building is embedded in all equilibrium strategies in the above literature. However, we have not seen works showing that cooperation and exploitation co-exist among homogeneous players.

7.2 Future Research

To be filled later.

Appendix A: On the Stationarity of Strategies

We demonstrate that if the matching pool strategy distribution is stationary, there will be an associated social state distribution. In turn, if this associated social state distribution is stationary, the original matching pool distribution remains stationary.

We should emphasize the causal direction is from matching pool distributuion to the social state distribution. In fact, it should be fairly obvious that, even if the social state distributuion happens to be stationary for some time, usually it would not create stationarity of matchig pool distribution.

Def. A matching state is a triple $\theta = (s, s'; t)$ where $s, s' \in S$ denote strategies used by two partners and $t \in Z := \{0, 1, 2, \dots\}$ denotes the number of periods that have passed since the partnership was created.

In particular, $(s, s'; 0)$ denotes a match which is just created by random pairing, and partners are of types s and s'. If $t > T(s, s')$, match must have been dissolved in $T(s, s')$ even if both partners survived and hence $(s, s'; t)$ should represent off-path state.

Note that the order of strategies is important. We treat that the state $\theta = (s, s'; t)$ is deffrent from the state $\theta' = (s', s; t)$ so that θ represents the state where s faces the t-th period of a partnership, while θ' represents the state where s' faces the t-th period of the same partnership.

Def. We denote by $P \in \Delta(S)$ the population distribution of strategies in the entire population. Thus, $P(s)$ denotes the population share (proportion of those players who adopt s) among the total population.

We denote by $Q(\theta) = Q(s, s'; t)$ the population share (proportion of those pairs who are in the state θ) among the total number of population. Note that the same pairing $(s, s'; t)$ is counted twice, once as $(s, s'; t)$ and another as $(s', s; t)$, making total number of matchings the same as the total population. We call Q as a **social state**.

$\sum_{s' \in S} Q(s, s'; 0) := q(s)$ is the share of strategy $s \in S$ in total population who are just matched with new partners, *i.e.,* who are in the matching pool at the beginning of the period. Hence, $q(s)$ represents the ratio of s strategies in the matching pool to the total population of the society. Finally, $\bar{q} := \sum_{s \in S}$ is the proportion of players who are in the matching pool (at the beginning of a period, but before matching takes place) in the entire population.

Obviously,

- $Q(s, s'; t) = 0$ if $t > T(s, s')$.
- $Q(s, s'; t) = 0$ if either $P(s) = 0$ or $P(s') = 0$ even if $t \leq T(s, s')$.
- $\sum_{s' \in S} \sum_{t \in Z} Q(s, s'; t) = P(s)$ for all $s \in S$.

We denote by $p \in \Delta(S)$ the starategy distribution in the matching pool. Thus, $p(s)$ is the ratio of strategy s among players in the matching pool.

By definition,

$$p(s) = \frac{q(s)}{\sum_{s' \in S} q(s')}.$$

Suppose matching pool distribution p and matching pool population ratio $\bar{p}$ are both stationary over periods. Then, the associated social state Q is defined as follows (see also the figure in the next page, which explains the dynamics of strategy c_1 population facing with various c_t strategies with $t \geq 1$.)

(a) First, $q(s)$ is constructed as $q(s) = \bar{q} \cdot p(s)$ for all $s \in S$.

(b) At the matching pool new partnerships are created. They are $Q(s, s'; 0) := q(s) \cdot q(s')$ for all $s, s' \in S$.

(c) If $T(s, s') \geq 1$, those partnerships which were created in the previous period or those partnerships that enter into $t = 1$ (in this period) must be

$$Q(s, s'; 1) := \delta^2 Q(s, s'; 0) = \delta^2 q(s) q(s')$$

because

- $(1 - \delta) Q(s, s'; 0)$ of type s players dies and are replaced by newly born s players at the beginning of this period, and
- $\delta(1 - \delta) Q(s, s'; 0)$ of type s players returns to the matching pool at the beginning of this period, because of her partner dies at the end of previous period, hence
- total of $(1 - \delta^2) Q(s, s'; 0)$ of the match state in $t = 0$ is dissolved at the end of the previous period, leaving $Q(s, s'; 1) := \delta^2 Q(s, s'; 0)$.

- Note that total of

$$Q_{mp}(s,s';1) = (1-\delta^2)Q(s,s';0) = (1-\delta^2)q(s)q(s')$$

type s players return to matchig pool from $Q(s,s';0)$ at the beginning of $t=1$.

(d) Those partnerships that enter into $t < T(s,s')$ (in this period) is

$$Q(s,s';t) = \delta^2 Q(s,s';t-1) = \delta^{2t}Q(s,s';0) = \delta^{2t}q(s)q(s').$$

Those effectively return to matching pool is

$$\begin{aligned} Q_{mp}(s,s';t) = (1-\delta^2)Q(s,s';t-1) &= (1-\delta^2)\delta^{2(t-1)}Q(s,s';0) \\ &= (1-\delta^2)\delta^{2(t-1)}q(s)q(s'). \end{aligned}$$

(e) If $t = T(s,s')$,

$$Q(s,s';T(s,s')) = \delta^2 Q(s,s';T(s,s')) = \delta^{2T(s,s')}q(s)q(s').$$

Those effectively return to matching pool is

$$Q_{mp}(s,s';T(s,s')) = Q(s,s';T(s,s')) = \delta^{2T(s,s')}q(s)q(s').$$

(f) If $t > T(s,s')$,

$$Q(s,s';t) = 0 \ \text{ and } Q_{mp}(s,s';t) = 0.$$

(g) In each period, $\sum_{t\in Z} Q_{pm}(s,s';t)$ returns to matching pool from matches with s' players, or

$$\begin{aligned} \sum_{t\in Z} Q_{pm}(s,s';t) &= \sum_{t=0}^{T(s,s')-1} (1-\delta^2)\delta^{2t}q(s)q(s') + \delta^{2T(s,s')}q(s)q(s') \\ &= (1-\delta^2)\frac{1-\delta^{2T(s,s')}}{1-\delta^2}q(s)q(s') + \delta^{2T(s,s')}q(s)q(s') \\ &= q(s)q(s') \end{aligned}$$

It follows that the total number (as a ratio of total population) of s players returning to matching pool is:

$$\sum_{s'\in S}\sum_{t\in Z} Q_{pm}(s,s';t) = \sum_{s'\in S} q(s)q(s') = q(s).$$

It follows that there are:

$$\sum_{t\in Z} Q(s,s';t) = \sum_{t=1}^{T(s,s')} \delta^{2t} q(s)q(s') = \frac{1-\delta^{2(T(s,s')+1)}}{1-\delta^2} q(s)q(s')$$

type s players who are in the various stages of matches with type s' players. Then, there are:

$$q(s) = \sum_{s'\in S}\sum_{t\in Z} Q(s,s';t) = \left[\sum_{s'\in S} \frac{1-\delta^{2(T(s,s')+1)}}{1-\delta^2} q(s')\right] q(s)$$

type s players in the total population.

Appendix B: Proofs

Proof of Lemma 2:

(a) Let s' be a strategy that chooses e in some t after on-path history. If $t \leq T$, 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

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

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 $t \leq T$ 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) Let s be any strategy that chooses D at some $t \geq T+1$ along on-path history.

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

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

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

□

Proof of Lemma 4:

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

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

If s' invades p, then for any $s \in supp(p)$,

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

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

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

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

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

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

Proof of Lemma 5:

From (1),

$$v(c_T; p_T^{T+1}(\alpha)) = \frac{\alpha r(c_T, c_T)}{r(c_T; p_T^{T+1}(\alpha))} v^I(c_T, c_T) + \frac{(1-\alpha) r(c_T, c_{T+1})}{r(c_T; p_T^{T+1}(\alpha))} v^I(c_T, c_{T+1}), \tag{12}$$

$$v(c_{T+1}; p_T^{T+1}(\alpha)) = \frac{\alpha r(c_{T+1}, c_T)}{r(c_{T+1}; p_T^{T+1}(\alpha))} v^I(c_{T+1}, c_T) + \frac{(1-\alpha) r(c_{T+1}, c_{T+1})}{r(c_{T+1}; p_T^{T+1}(\alpha))} v^I(c_{T+1}, c_{T+1}). \tag{13}$$

By computation

$$\begin{aligned}
r(c_T, c_T) &= (1-\delta)\{1+\delta^2+\cdots\} = \frac{1}{1+\delta} = r(c_{T+1}, c_{T+1}); \\
r(c_T, c_{T+1}) &= (1-\delta)\{1+\delta^2+\cdots+\delta^{2T}\} = \frac{1-\delta^{2(T+1)}}{1+\delta} = r(c_{T+1}, c_T); \\
r(c_T; p_T^{T+1}(\alpha)) &= \frac{\alpha}{1+\delta} + \frac{(1-\alpha)(1-\delta^{2(T+1)})}{1+\delta}; \\
r(c_{T+1}; p_T^{T+1}(\alpha)) &= \frac{\alpha(1-\delta^{2(T+1)})}{1+\delta} + \frac{(1-\alpha)}{1+\delta};
\end{aligned}$$

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

Therefore

$$v(c_T; p_T^{T+1}(\alpha)) = \frac{(1-\delta^{2T})d + \delta^{2T}(1-\delta^2)\ell + \delta^{2T}[c - (1-\delta^2)\ell]\alpha}{1-\delta^{2(T+1)} + \alpha\delta^{2(T+1)}}. \tag{14}$$

$$v(c_{T+1}; p_T^{T+1}(\alpha)) = \frac{(1-\delta^{2(T+1)})d + \delta^{2(T+1)}c + \delta^{2T}[(1-\delta^2)(g-d) - \delta^2 c]\alpha}{1-\alpha\delta^{2(T+1)}}. \tag{15}$$

By computation

$$\begin{aligned}
& \frac{(1-\delta^{2(T+1)}+\alpha\delta^{2(T+1)})(1-\alpha\delta^{2(T+1)})}{\delta^{2T}}\Big\{v(c_T;p_T^{T+1}(\alpha))-v(c_{T+1};p_T^{T+1}(\alpha))\Big\} \\
= & \frac{1}{\delta^{2T}}\Big[(1-\alpha\delta^{2(T+1)})\{(1-\delta^{2T})d+\delta^{2T}(1-\delta^2)\ell+\delta^{2T}[c-(1-\delta^2)\ell]\alpha\} \\
& \quad +(1-\delta^{2(T+1)}+\alpha\delta^{2(T+1)})\{(1-\delta^{2(T+1)})d+\delta^{2(T+1)}c+\delta^{2T}[(1-\delta^2)(g-d)-\delta^2c]\alpha\}\Big] \\
= & \frac{1}{\delta^{2T}}\Big[(1-\delta^{2T})d+\delta^{2T}(1-\delta^2)\ell-(1-\delta^{2(T+1)})\{(1-\delta^{2(T+1)})d+\delta^{2(T+1)}c\} \\
& -\alpha\delta^{2(T+1)}\{(1-\delta^{2T})d+\delta^{2T}(1-\delta^2)\ell\}+\alpha\delta^{2T}\{c-(1-\delta^2)\ell\} \\
& -\alpha\delta^{2(T+1)}\{(1-\delta^{2(T+1)})d+\delta^{2(T+1)}c\}-\alpha(1-\delta^{2(T+1)})\delta^{2T}\{(1-\delta^2)(g-d)-\delta^2c\} \\
& \alpha^2\delta^{4T+2}\{c-(1-\delta^2)\ell\}-\alpha^2\delta^{4T+2}\{(1-\delta^2)(g-d)-\delta^2c\}\Big] \\
= & -\{(1-\delta^{2(T+1)})(c-d)+(1-\delta^2)(d-\ell)\} \\
& +\alpha[\{1+\delta^2-2\delta^{2T+4}\}c+\{1-3\delta^2+2\delta^{2T+4}\}d-(1-\delta^2)\{(1+\delta^{2(T+1)})\ell+(1-\delta^{2(T+1)})g\}] \\
& -\alpha^2\delta^{2(T+1)}(1-\delta^2)(c-\ell+g-d) \\
= & \Delta\tilde{v_T}(\alpha,\delta).
\end{aligned}$$

Therefore,

$$\Delta\tilde{v_T}(\alpha,\delta) := A_T(\delta)\alpha^2+B_T(\delta)\alpha+C_T(\delta), \tag{16}$$

where

$$\begin{aligned}
A_T(\delta) &= -\delta^{2(T+1)}(1-\delta^2)(c-\ell+g-d)<0; \\
B_T(\delta) &= \{1+\delta^2-2\delta^{2T+4}\}c+\{1-3\delta^2+2\delta^{2T+4}\}d-(1-\delta^2)\{(1+\delta^{2(T+1)})\ell+(1-\delta^{2(T+1)})g\}; \\
C_T(\delta) &= -\{(1-\delta^{2(T+1)})\delta^2(c-d)+(1-\delta^2)(d-\ell)\}<0.
\end{aligned}$$

The claims in the lemma follow immediately. □

Proof of Lemma 6:

Let $\mu_T(\alpha)=\frac{\alpha r(c_T,c_T)}{r(c_T;p_T^{T+1}(\alpha))}$ and $\mu_{T+1}(\alpha)=\frac{\alpha r(c_{T+1},c_T)}{r(c_{T+1};p_T^{T+1}(\alpha))}$. Then (12) and (13) 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,

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

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,

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

□

Proof of Lemma 7:

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 continuation payoff of c_t-strategy at the beginning of $t+1$ is

$$
V(c_t; p, t+1) = \alpha_t \{ \frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2} V(c_t; p) \} + (1-\alpha_t) \{ \ell + \delta V(c_t; p) \}. \tag{17}
$$

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

$$
\begin{aligned}
V(c_{t+1}; p, t+1) &= \alpha_t \{ g + \delta V(c_{t+1}; p) \} \\
&\quad + (1-\alpha_t) \{ d + \delta(1-\delta) V(c_{t+1}; p) + \delta^2 V(c_{t+1}; p, t+2) \}.
\end{aligned} \tag{18}
$$

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 (18) becomes

$$\begin{aligned} V(c_{t+1}; p, t+1) &= \alpha_t\{g + \delta V(c_{t+1}; p)\} \\ &\quad +(1-\alpha_t)\{d + \delta(1-\delta)V(c_{t+1}; p) + \delta^2 V(c_t; p, t+1)\} \\ \Longleftrightarrow V(c_{t+1}; p, t+1) &= \frac{\alpha_t\{g + \delta V(c_{t+1}; p)\} + (1-\alpha_t)\{d + \delta(1-\delta)V(c_{t+1}; p)\}}{1-(1-\alpha_t)\delta^2}. \end{aligned} \tag{19}$$

From the assumption,

$$V(c_t; p) = V(c_{t+1}; p). \tag{20}$$

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

$$V(c_t; p, t+1) = V(c_{t+1}; p, t+1). \tag{21}$$

(21) implies that the RHS of (17) and (19) must be the same. Using (20) 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 (20) implies $\alpha_t = \alpha$ for all $t = T, T+1, \ldots$. This implies that the fraction of $c_{T+\tau}$-strategy is of the form $\alpha(1-\alpha)^\tau$. □

Proof of Lemma 8:

Under the fraction structure of $p_T^\infty(\alpha)$, the payoff sequence of strategies are as follows.

Table 3(a)-(c) show the sequence of payoffs for c_T, c_{T+1}, and c_{T+2} within a match.

$p_T^\infty(s')$	$s' \setminus$ time	1	2	$\cdots$	$T-1$	T	$T+1$	$T+2$	$T+3$
α	c_T	d	d	$\cdots$	d	c	c	c	$\cdots$
$(1-\alpha)$	c_{T+1} and up	d	d	$\cdots$	d	ℓ			

Table 3(a): Payoff sequence of c_T-strategy under $p_T^\infty(\alpha)$ within a match against s'.

$p_T^\infty(s')$	$s' \setminus$ time	1	2	$\cdots$	$T-1$	T	$T+1$	$T+2$	$T+3$
α	c_T	d	d	$\cdots$	d	g			
$(1-\alpha)\alpha$	c_{T+1}	d	$\mathbf{d}$	$\cdots$	$\mathbf{d}$	$\mathbf{d}$	$\mathbf{c}$	$\mathbf{c}$	$\cdots$
$(1-\alpha)^2$	c_{T+2} and up	d	$\mathbf{d}$	$\cdots$	$\mathbf{d}$	$\mathbf{d}$	ℓ		

Table 3(b): Payoff sequence of c_{T+1}-strategy under $p_T^\infty(\alpha)$ within a match against s'.

$p_T^\infty(s')$	$s' \setminus$ time	1	2	$\cdots$	$T-1$	T	$T+1$	$T+2$	$T+3$
α	c_T	d	d	$\cdots$	d	g			
$(1-\alpha)\alpha$	c_{T+1}	d	$\mathbf{d}$	$\cdots$	$\mathbf{d}$	$\mathbf{d}$	$\mathbf{g}$		
$(1-\alpha)^2\alpha$	c_{T+2}	d	$\mathbf{d}$	$\cdots$	$\mathbf{d}$	$\mathbf{d}$	$\mathbf{d}$	$\mathbf{c}$	$\cdots$
$(1-\alpha)^3$	c_{T+3} and up	d	$\mathbf{d}$	$\cdots$	$\mathbf{d}$	$\mathbf{d}$	$\mathbf{d}$	ℓ	

Table 3(c): Payoff sequence of c_{T+2}-strategy under $p_T^\infty(\alpha)$ within a match against s'.

Notice that the bold-faced sub-table of Table 3(b) is identical to the Table 3(a). This is because from 2nd period on, c_{T+1}-strategy behaves the same way as c_T-strategy against itself and $c_{T+2},\ldots$ strategies.

Similarly, after 2nd period on, c_{T+2}-strategy behaves the same way as c_{T+1}-strategy against itself and $c_{T+3},\ldots$ strategies.

In equation, the long-run payoff of c_T-strategy is decomposed as

$$\begin{aligned} V(c_T; p_T^\infty(\alpha)) &= \alpha V(c_T, c_T; p_T^\infty(\alpha)) \\ &\quad +(1-\alpha)V(c_T, c_{T+1}; p_T^\infty(\alpha)). \end{aligned} \tag{22}$$

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

$$\begin{aligned} V(c_{T+1}; p_T^\infty(\alpha)) &= \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) \\ &\quad +(1-\alpha)[\alpha\{d + \delta^2 V(c_T, c_T; p_T^\infty(\alpha)) + \delta(1-\delta)V(c_{T+1}; p_T^\infty(\alpha))\} \\ &\qquad (1-\alpha)\{d + \delta^2 V(c_T, c_{T+1}; p_T^\infty(\alpha)) + \delta(1-\delta)V(c_{T+1}; p_T^\infty(\alpha))\}] \\ &= \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) \\ &\quad +(1-\alpha)[d + \delta^2 V(c_T; p_T^\infty(\alpha)) + \delta(1-\delta)V(c_{T+1}; p_T^\infty(\alpha))], \end{aligned} \tag{23}$$

where the last equality uses (22).

Equivalently we can write

$$[1-(1-\alpha)\delta(1-\delta)]V(c_{T+1}; p_T^\infty(\alpha)) = \alpha V(c_{T+1}, c_T; p_T^\infty(\alpha)) + (1-\alpha)d + (1-\alpha)\delta^2 V(c_T; p_T^\infty(\alpha)). \tag{24}$$

Similarly from Table 3(b) and 3(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

$$[1-(1-\alpha)\delta(1-\delta)]V(c_{T+2};p_T^{\infty}(\alpha)) = \alpha V(c_{T+1},c_T;p_T^{\infty}(\alpha))+(1-\alpha)d+(1-\alpha)\delta^2 V(c_{T+1};p_T^{\infty}(\alpha)). \quad (25)$$

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 (4) and (5) 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 $t > T$. □

Proof of Lemma 9:

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,

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

(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 4.)

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 $\alpha = 1$. □

Proof of lemma 10:

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,

$$
v(c_{T+\tau}; p_T^\infty(\alpha)) = \frac{w + \alpha(1-\alpha)^\tau r(c_{T+\tau}, c_{T+\tau})v^I(c_{T+\tau}, c_{T+\tau}) + (1-\alpha)^{\tau+1} r(c_{T+\tau}, c_{T+\tau+1})v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1 + \alpha(1-\alpha)^\tau r(c_{T+\tau}, c_{T+\tau}) + (1-\alpha)^{\tau+1} r(c_{T+\tau}, c_{T+\tau+1})}
$$

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

$$
v(s_\tau; p_T^\infty(\alpha)) = \frac{w + \alpha(1-\alpha)^\tau r(s_\tau, c_{T+\tau})v^I(s_\tau, c_{T+\tau}) + (1-\alpha)^{\tau+1} r(c_{T+\tau}, c_{T+\tau+1})v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1 + \alpha(1-\alpha)^\tau r(s_\tau, c_{T+\tau}) + (1-\alpha)^{\tau+1} r(c_{T+\tau}, c_{T+\tau+1})}.
$$

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

$$
\begin{aligned}
v(c_{T+\tau}; p_T^\infty(\alpha)) &= \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big)\frac{R_1+R_3}{R_1+R_2+R_3} + \frac{R_2 v^I(c_{T+\tau}, c_{T+\tau})}{R_1+R_2+R_3} \\
&= \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big) \\
&\quad + \frac{R_2}{R_1+R_2+R_3}\Big[v^I(c_{T+\tau}, c_{T+\tau}) - \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big)\Big]. \quad (26)
\end{aligned}
$$

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

$$
\begin{aligned}
v(s_\tau, ; p_T^\infty(\alpha)) &= \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big)\frac{R_1+R_3}{R_1+R_2'+R_3} + \frac{R_2' v^I(s_\tau, c_{T+\tau})}{R_1+R_2'+R_3} \\
&= \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big) \\
&\quad + \frac{R_2'}{R_1+R_2'+R_3}\Big[v^I(s_\tau, c_{T+\tau}) - \Big(\frac{w}{R_1+R_3} + \frac{R_3 v^I(c_{T+\tau}, c_{T+\tau+1})}{R_1+R_3}\Big)\Big]. \quad (27)
\end{aligned}
$$

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 (26) is larger than that of (27). □

Proof of Proposition 3:

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 $T \le \tau < \infty$. 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 $t \ge T+1$: 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 $\tau \ge T$: 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 10.

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.

□

Proof of Lemma 12:

any NE strategy of VRPD, $p' \in \mathcal{P}(\boldsymbol{S})$, and define any assocaited babbling strategy $p\sigma^B(p) \in \mathcal{P}(\boldsymbol{S}^{CT})$. If there is another strategy $p \in \mathcal{P}(\boldsymbol{S})$, whose expected payoff is higher than that of p, then the following neologism strategy ("early harvest" neologism strategy), $\sigma^N(p;p') \in \boldsymbol{S}^{CT}$, such that it invades p^B with the following features.

1. $\sigma^N(p;p')$ uses a neologism, i.e., it uses a message which is not being used by any strategy in supp$(NE(p'))$,

2. $\sigma^N(p;p') = p$, i.e., when they receives the neologism, it plays supp$p$ or plays $s \in$ supp(p) with probability $p(s)$,

3. $\sigma^N(p;p') = p'$, i.e., when they receives non-neologism messages, it plays suppp' or plays $s \in$ supp(p') with probability $p'(s)$.

Clealy $\sigma^N(p;p')$ can invade p'^B. However, $NE(p)$ to be an EES, p must be a best response against the post-entry distribution, $(1-\varepsilon p') + \varepsilon p$. Unless p itself is a NSS, such property does not hold. Hence the assertion holds. □

References

Bowles, A. and R. Boyer (1988) "Labor Discipline and Aggregate Demand: A Macroeconomic Model," *American Economic Review*, 78(2), 395-400.

Carmichael L. and B. MacLeod (1997) "Gift Giving and the Evolution of Cooperation," *International Economic Review*, 38(3), 485-509.

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

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(4), 533-556.

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

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

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

Kranton, R. (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.

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

Matsui, A., and Okuno-Fujiwara, M. (2002) "Evolution and the Interaction of Conventions," *The Japanese Economic Review*, 53(2), 141-153.

Okuno-Fujiwara, M. (1987) "Monitoring Cost, Agency Relationship, and Equilibrium Modes of Labor Contract," *Journal of the Japanese and International Economies*, 1(2), 147-167.

Okuno-Fujiwara, M., and Postlewaite A. (1995). "Social Norms and Random Matching Games," *Games and Econonomic Behavior*, 9(1), 79-109.

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(3), 433-444.

Swinkels, J (1992) "Evolutionary Stability with Equilibrium Entrants," *Journal of Economic Theory*, 57(2), 306-332.