CIRJE-F-599

Voluntarily Separable Repeated Prisoner's Dilemma

Takako Fujiwara-Greve
Keio University

Masahiro Okuno-Fujiwara
University of Tokyo

October 2008

Voluntarily Separable Repeated Prisoner's Dilemma*

Takako Fujiwara-Greve   Masahiro Okuno-Fujiwara

Keio University†   University of Tokyo‡

First Draft: August, 2006
This Version: May 25, 2008

Abstract

Ordinary repeated games do not apply to real societies where one can cheat and run away from partners. We formulate a model of endogenous relationships which one can unilaterally end and start with a randomly-assigned new partner with no information flow. Focusing on two-person, two-action Prisoner's Dilemma, we show that the endogenous duration of partnerships gives rise to a significantly different evolutionary stability structure from ordinary random matching games. *Monomorphic* equilibria require initial trust-building, while a *polymorphic* equilibrium includes early cooperators than any strategy in monomorphic equilibria and is thus more efficient. This is due to nonlinearity of average payoffs.

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

---

*We are very grateful to Nobue Suzuki for her constant help. We also thank useful comments from the editor, Juuso Välimäki, two anonymous referees, Akihiko Matsui, Chiaki Hara, Masahiro Ashiya, Mamoru Kaneko, Masaki Aoyagi, Marcus Berliant, and seminar participants at KIER, Kyoto University, Keio University, University of Tokyo, University of Tsukuba, Hitotsubashi University, and Symposium on Market Quality (Keio University, Tokyo, December 2005). All errors are ours. This work was supported by Grant-in-Aid for Scientific Research on Priority Areas, No.12124204.

†Department of Economics, 2-15-45 Mita, Minatoku, Tokyo 108-8345 JAPAN. E-mail: takakofg@ econ.keio.ac.jp

‡Graduate School of Economics, 7-3-1 Hongo, Bunkyoku, Tokyo 113-0033 JAPAN. E-mail: mfujiwar@e.u-tokyo.ac.jp

1 Introduction

Economic transactions are often threatened by moral hazard. Historically, many devices are made to mitigate moral hazard, from face-to-face transactions to legal enforcement of contracts and rights. As the transactions have become global and the age of Internet has come, such traditional systems of moral enforcements may become less effective, because of the anonymity of trade partners and the ease of entry to and exit from transactions. We believe, however, that such social change does not destroy beneficial transactions but simply requires other forms of moral enforcement. In this paper we analyze an evolutionary model of endogenous long-term relationships with no information flow beyond the current partner and show a hitherto not analyzed way to discipline players to cooperate.

We consider the following model. There is a large society of homogeneous players, who are randomly matched to play a component game if they have no partner at the beginning of a period. The component game is a repeated Prisoner's Dilemma until either partner wants to break up. We call this component game as "Voluntarily Separable Repeated Prisoner's Dilemma (VSRPD)". If break-up occurs, each player finds a new partner by random match and starts another VSRPD. When a pair of players start a new relationship, it is assumed that neither player knows the past actions of the other. Therefore, a strategy is how to play a VSRPD, starting from a null history each time with a new partner. Many transactional relationships fit our model. Workers can shirk, quit, and find a new employer without telling the past, and borrowers can move to another city after defaulting and find a new lender without telling the true credit history.

In order to make a cooperative relationship, it is necessary to reduce the continuation payoff of a defector. Ordinary trigger strategies do not work because the punishment cannot be applied to a defector who can run away and restart with a new partner with no information flow. After a break-up, the lifetime payoff starting from a new partnership is the continuation payoff, and thus it must not be too large to deter defection. In the literature, matching friction (unemployment) and gradual cooperation (trust building) have been introduced to lower the life-time payoff and analyzed extensively.[1] These mechanisms impose low payoff after a break-up or reduce initial payoff of new matches and sustain symmetric (*monomorphic*) cooperative equilibria.

We show that there is a third disciplining device: co-existence of diverse strategies. If there are different types of strategies in the population and if it is beneficial to match with the same type but not with a different type, such *polymorphic* distribution sustains cooperation among partnerships with the same type, because defection and break-up lead to a possible bad match.

[1]See Shapiro and Stiglitz (1984), and Okuno-Fujiwara (1987) for unemployment as a disciplining device and Carmichael and MacLeod (1997), Datta (1996), Ghosh and Ray (1996), Kranton (1996a) and Watson (2002) for gradual cooperation.

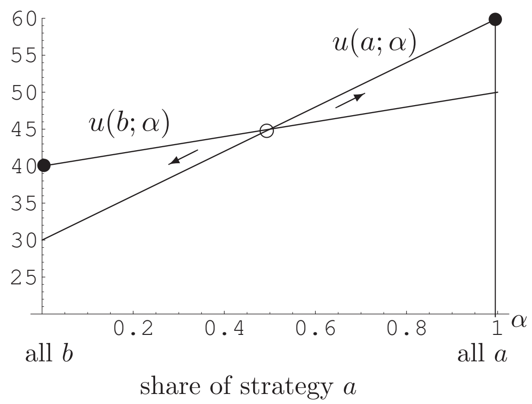

Figure 1a: Ordinary random matching

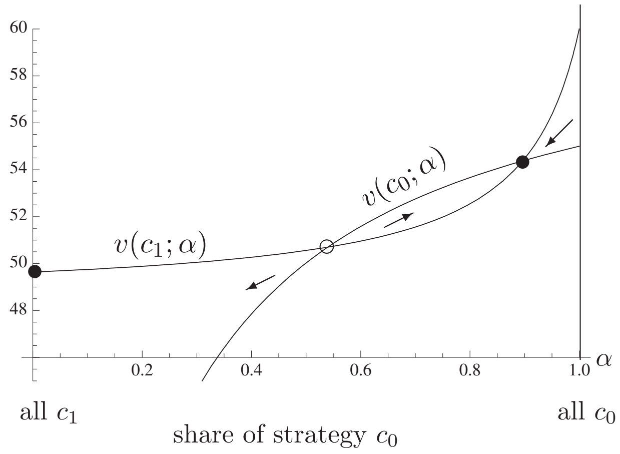

Figure 1b: Endogenous partnerships

Moreover, a polymorphic equilibrium is more efficient than any monomorphic equilibrium, since the latter requires that everyone must build trust initially, while the former can include strategies that start cooperating earlier. This is a substantially different conclusion compared to those of the random matching games.

The existence and payoff superiority of a stable polymorphic equilibrium are thanks to the nonlinearity of payoffs under endogenous length of the VSRPD. In the usual random matching games, the component game's length is exogenously fixed, and therefore the payoff of a strategy is linear in the strategy-share distribution. Consider, for example, a random matching game of a one-shot 2 by 2 coordination game. Let a and b be strategies and α be the fraction of a-strategy in the population. The expected payoff of x-strategy ($x = a, b$) is linear in α such that $u(x;\alpha) = \alpha u(x,a) + (1-\alpha)u(x,b)$. A numerical example[2] is displayed in Figure 1a, and we see that the mixed-strategy equilibrium is unstable and less efficient than (at least one) efficient pure equilibrium.

In our model, however, the duration of VSRPD is endogenous. If there are different strategies in the population and only a match with the same strategy would last, then the lifetime average payoff becomes nonlinear, which can be intuitively explained as follows. Let c_0 and c_1 be strategies and α be the fraction of c_0-strategy in the population. Assume that if the same strategies are matched, the partnership continues as long as they both live (with probability δ^2 in each period), and a c_0-pair earns higher payoff within the match than a c_1-pair. However, if different strategies are matched, c_1-strategy exploits c_0-strategy and the partnership ends immediately. Let the per period average payoff of c_t strategy within a match with $c_{t'}$-strategy be $v(c_t, c_{t'})$.

The lifetime average payoff of each strategy before knowing the new partner is the total

[2] Based on $u(a,a) = 60 > u(b,b) = 40$, $u(a,b) = 30$, and $u(b,a) = 50$.

expected payoff divided by the total expected duration of matches;[3]

$$v(c_0;\alpha) = \frac{\alpha \frac{1}{1-\delta^2} v(c_0,c_0) + (1-\alpha)\cdot 1 \cdot v(c_0,c_1)}{\alpha \frac{1}{1-\delta^2} + (1-\alpha)\cdot 1},$$

$$v(c_1;\alpha) = \frac{\alpha \cdot 1 \cdot v(c_1,c_0) + (1-\alpha) \frac{1}{1-\delta^2} v(c_1,c_1)}{\alpha \cdot 1 + (1-\alpha)\frac{1}{1-\delta^2}}.$$

Therefore they are non-linear in α. A numerical example is depicted in Figure 1b.

Moreover, we can show that the "victim" c_0-strategy has a concave lifetime average payoff and the exploiter c_1-strategy has a convex lifetime average payoff as follows. Consider the effect of **decrease** of α from 1 to 0 on the average payoff of c_0-strategy. As the share of c_1-strategy increases, c_0-strategy gets exploited more often, and thus its payoff decreases. Moreover, this payoff decrease accelerates as α decreases, because (i) partnership dissolutions due to mismatch occur more often, (ii) thus c_0-strategy goes into the matching pool more often, and (iii) hence c_0-strategy meets c_1-strategy more often. Therefore the average payoff of c_0-strategy is concave in α, as in Figure 1b.

Similarly, as α **increases** from 0 to 1, the average payoff of c_1-strategy increases because the share of victims increases. Moreover, the payoff is convex in α, since the matching rate with the victims in its lifetime increases more as α increases, thanks to more frequent dissolution of partnerships. In this way, the endogenous duration of partnerships implies convex/concave payoff functions, which was not the case in ordinary random matching games.

The nonlinear payoff structure gives rise to a substantially different stability of strategy distributions from the ordinary random matching games. In Figure 1b, there are two stable equilibria; the monomorphic equilibrium of c_1 ($\alpha = 0$) and the *bimorphic* equilibrium where $\alpha \approx 0.88$. The middle Nash equilibrium where $\alpha \approx 0.5$ is not stable. Furthermore, since both payoff functions are increasing in the share of c_0-strategy, the bimorphic equilibrium with a large share of c_0-strategy is more efficient than the most efficient monomorphic equilibrium, consisting only of c_1-strategy. (The monomorphic distribution of c_0-strategy where $\alpha = 1$ does not constitute even a Nash equilibrium.)

This evolutionary stability of polymorphic distributions in a homogeneous population provides a foundation of incomplete information models such as Ghosh and Ray (1997) and Rob and Yang (2005), because it illustrates that different "types" of strategies may appear spontaneously through evolutionary processes even if all players have the same characteristics (stage game payoff function, information structure, and matching probabilities). In a bargaining context, Abreu and Sethi (2003) have a similar motivation to ours to endogenize multiple behavior rules through evolution. Their focus is to derive co-existence of rational and behavioral players, while we give co-existence of different strategies among homogeneous players.

[3]See Section 2.2 for details.

P1 \ P2	C	D
C	c, c	ℓ, g
D	g, ℓ	d, d

Table 1: Payoff of Prisoner's Dilemma

This paper is organized as follows. In Section 2, we introduce the formal model and stability concepts. In Section 3, we show various neutrally stable strategy distributions and that more diversity implies higher equilibrium payoff. In Section 4 we discuss extensions, and in Section 5 we give concluding remarks.

2 Model and Stability Concepts

2.1 Model

Consider a society with a continuum of players. Each player may die in each period $1, 2, \ldots$ with probability $0 < (1-\delta) < 1$. When players die, they are replaced by newborn players, keeping the total population constant. Newborn players and players who do not have a partner enter into the matching pool where players are randomly paired to play the following *Voluntarily Separable Repeated Prisoner's Dilemma (VSRPD).*

In each period, matched players play ordinary two-person, two-action Prisoner's Dilemma, whose actions are denoted as Cooperate and Defect. After observing the play action profile of the period, they choose simultaneously whether or not they want to keep the match into the next period (action k) or bring it to an end (action e). Unless both choose k, the match is dissolved and players will have to start the next period in the matching pool. In addition, even if they both choose k, one's partner may die with probability $1-\delta$ which forces the surviving 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 partners play Prisoner's Dilemma again.

Assume that there is limited information available to play VSRPD. In each period, players know the 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 by the partners determines their one-shot payoffs. We denote the payoffs associated with each play action profile as[4]: $u(C,C) = c$, $u(C,D) = \ell$, $u(D,C) = g$, $u(D,D) = d$ with the ordering $g > c > d > \ell$ and[5] $2c \geqq g + \ell$. (See Table 1.)

We assume that the innate discount rate is zero except for the possibility of death, hence

[4]The first coordinate is your own action.

[5]The assumption $2c \geqq g + \ell$ is to simplify the analysis by making the symmetric action profile (C,C) most efficient. It is possible to construct equilibria to play (C,D) and (D,C) alternatingly when $2c < g + \ell$. See Section 4.2.

each player finds the relevant discount factor to be $\delta \in (0,1)$. With this, lifetime payoff for each player is well-defined given his own strategy (for VSRPD) 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 no-information-flow assumption, we focus on match-independent strategies that only depend on t and the private history of actions in the Prisoner's Dilemma within a match.[6] Let $H_t := \{C, D\}^{2(t-1)}$ be *the set of partnership histories at the beginning of* $t \geqq 2$ and let $H_1 := \{\emptyset\}$.

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

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

We investigate the evolutionary stability of *stationary* strategy distributions in the matching pool. Although the strategy distribution in the matching pool may be different from the distribution in the entire society, if the former is stationary, the distribution of various states of matches is also stationary, thanks to the stationary death process.[7]

In addition to the stability of social states, stationarity in the matching pool allows us to compute the lifetime payoff of each strategy explicitly, because it becomes recursive. (See equation (1) in the next subsection.)

2.2 Lifetime and Average Payoff

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

[6] The continuation decision is observable, but strategies cannot vary depending on combinations of $\{k, e\}$ since only (k, k) will lead to the future choice of actions. It is also possible to allow strategies to depend on a player's personal history, like Kandori (1992) and Ellison (1994). However, the cooperation using "contagion of defection" in their models does not hold in our model of continuum of players.

[7] To explain, let a *social state* be the distribution of strategy pairs in the society classified according to the "planned length" of the pair, i.e., when the pair intends to end the partnership if no death occurs, in a VSRPD. A social state is determined by in-flows from matching pool and out-flows by the random death and the time that passes. If the strategy distribution in the matching pool is stationary, the distribution of newly formed strategy pairs are stationary. Each newly formed pair has a planned length of the partnership (possibly infinite) determined by their strategies. Let $n = 1, 2, 3, \ldots, \infty$ be all possible planned lengths of strategy pairs. Since the random death applies to all pairs with the same probability in each period and the planned length of a partnership becomes one period less for each pair after each period (except for $n = \infty$), the outflow from each group of pairs with the same n is stationary, and newly formed pairs will join group n's in a stationary manner. Therefore the distribution of groups of pairs with the same n is stationary as well.

maintain the match, it will only continue with probability δ^2. Suppose that the planned length of the partnership of s and s' is $T(s,s')$ periods, if no death occurs. Then

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

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

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

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

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

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

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

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

where $supp(p)$ is the support of the distribution p, the sum $\delta(1-\delta)\{1+\delta^2+\cdots+\delta^{2\{T(s,s')-2\}}\}$ is the probability that s loses the partner s' before $T(s,s')$, and $\delta^{2\{T(s,s')-1\}}\delta$ is the probability that the match continued until $T(s,s')$ and s survives at the end of $T(s,s')$ and goes back to the matching pool. Note that the stationarity of p implies that the continuation payoff is always $V(s;p)$ after a match ends for any reason.

Let $L(s;p) := \sum_{s' \in supp(p)} p(s')L(s,s')$. By computation, (1) becomes

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

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

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

where the ratio $L(s,s')/L(s;p)$ is the relative length of periods that s expects to spend in a match with s'. As noted in the Introduction, this nonlinearity is due to the endogenous duration of partnerships. Note also that, if p is a strategy distribution consisting of a single strategy s', then $v(s;p) = v(s,s')$.

2.3 Nash Equilibrium

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

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

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

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

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

Proof: Consider a myopic strategy $\tilde{d}$ which plays D at $t=1$ and ends the partnership for any observation at $t=1$. For $t \geqq 2$, which is off-path, specify arbitrary actions. Then any $\tilde{d}$-strategy earns g as the average payoff under p, which is the maximal possible payoff. I.e., $\tilde{d} \in BR(p)$ and $s \notin BR(p)$ for all $s \in \text{supp}(p)$. □

Therefore, the C-trigger and C-tit-for-tat strategy of ordinary repeated Prisoner's Dilemma cannot constitute even a Nash equilibrium. Some fraction of players must play D in the first period of a partnership, in any equilibrium.

By contrast, $p_{\tilde{d}}$ consisting only of a "hit-and-run" $\tilde{d}$-strategy is a Nash equilibrium. Against a $\tilde{d}$-strategy, any strategy must play one-shot Prisoner's Dilemma. Hence, any strategy that starts with C in $t=1$ earns strictly lower average payoff than that of a $\tilde{d}$-strategy, and any strategy that starts with D in $t=1$ earns the same average payoff as that of a $\tilde{d}$-strategy.

2.4 Neutral Stability

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

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

Definition: A stationary strategy distribution $p \in \mathcal{P}(\boldsymbol{S})$ in the matching pool is a *Neutrally Stable Distribution* (NSD) if, for any $s' \in \boldsymbol{S}$, there exists $\bar{\epsilon} \in (0,1)$ such that for any $s \in supp(p)$ and any $\epsilon \in (0, \bar{\epsilon})$,

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

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

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

A stronger notion of stability that requires strict inequality (which is used in the notion of Evolutionary Stable Strategy) is too strong in our extensive-form model since any strategy that is different in the off-path actions from the incumbent strategies can earn the same average payoff as the incumbents'.

Similar to the ordinary "static" notion of evolutionary stability, our definition is based on the assumption that mutation takes place rarely so that only single mutation occurs within the time span in which a stationary strategy distribution is formed. However, unlike the ordinary notion of neutral stability (or ESS) of one-shot games, we need to assume the expected length of the lifetime of a mutant strategy in order to calculate its average payoff. In ordinary evolutionary games, the length of the component game is exogenously fixed, and so is the length of the lifetime of a mutant. In our model, by contrast, the length of the partnership is endogenous, and thus there is no obvious way to define the lifetime of a mutant. Since the partnership can potentially continue forever, we required that a stationary distribution of a mutant strategy to be deterred. This is the strongest notion of stability, because any shorter-lived mutants can be deterred. While we do not insist that the above definition is the best imaginable, it is tractable and justifiable.

We show that any myopic $\tilde{d}$-strategy is not a NSS, even though it constitutes a monomorphic Nash equilibrium.[8] Hence NSD concept selects among Nash equilibria in our model.

Lemma 2. *Any myopic $\tilde{d}$-strategy is not a NSS.*

Proof: Consider the following c_1-strategy.

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

[8] Note that repeated Defection itself can be sustained by some NSS, as later analysis shows. For example, a strategy that always defects but keeps the partnership if and only if (D, D) is observed is a NSS.

For any $\epsilon \in (0,1)$, let $p := (1-\epsilon)p_{\tilde{d}} + \epsilon p_{c1}$. From (3),

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

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

2.5 Trust-building Strategies

The successful invader c_1-strategy generates the most efficient symmetric outcome among those that play D at least once. However, it may not constitute a symmetric Nash equilibrium if the deviation payoff g is too large and the survival rate δ is too small. We thus focus on its generalized versions called *trust-building strategies*, defined below.

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

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

The first T periods of c_T-strategy are called *trust-building periods* and the periods afterwards are called *cooperation periods*. A trust-building strategy continues the partnership if and only if "acceptable" action profiles are played, and the acceptable action profile during the trust-building periods is (D,D) only and during the cooperation periods is (C,C) only.

In this paper we are not trying to establish a folk theorem but instead we investigate how much efficiency can be attained. Playing C forever after some point is desirable for efficiency. Since a player can unilaterally end the partnership, ending the partnership is the maximal punishment. Thus, the trust-building strategies are sufficient to look for the second best.

Needless to say, Nash equilibrium and NSD are proved by checking all other strategies in $\boldsymbol{S}$ (not just among trust-building strategies).

3 Neutrally Stable Distributions

3.1 Monomorphic NSS

We first consider monomorphic strategy distributions, consisting of a single c_T-strategy, as a benchmark. The literature of endogenous partnerships has focused on symmetric strategy distributions. In the literature, multiple-action Prisoner's Dilemma was often used and thus gradual increase of cooperation level was feasible. In our model, there are only two actions and thus c_T-strategy can be interpreted as a "gradual cooperation" strategy.

Let p_T be the strategy distribution consisting only of c_T-strategy. We first derive a condition on T to warrant that p_T is a Nash equilibrium. By the usual logic of dynamic programming, it suffices to prove that the average payoff generated by one-shot deviation is not higher than that of c_T-strategy (i.e., use of a strategy that differ from c_T in one-step in a VSRPD, followed by c_T-strategy from the next VSRPD on, do not fare better than c_T-strategy). Note that in our model, there are two phases that one can deviate, in the Prisoner's Dilemma and in the continuation decision phase.

It is straightforward to show that one-shot deviations to end the partnership after observing on-path actions and one-shot deviations to play C during the trust-building periods do not earn higher payoff than c_T does. Therefore it is sufficient to deter one-shot deviation strategies that play D when $t \geqq T+1$. During the cooperation periods, a one-shot deviation strategy earns g immediately but then goes back to the matching pool if he survives. Therefore, the (non-averaged) continuation payoff is

$$g + \delta V(c_T; p_T) = g + (L-1)v(c_T; p_T).$$

By contrast, c_T receives one-shot payoff c as long as the partners survive but goes back to the matching pool if he outlives the partner. From (2), the continuation payoff is

$$\frac{c}{1-\delta^2} + \{1 - \frac{L(c_T, c_T)}{L}\}V(c_T; p_T) = \frac{c}{1-\delta^2} + \{L - \frac{1}{1-\delta^2}\}v(c_T; p_T).$$

Hence one-shot deviations during the cooperation periods of a partnership is deterred if

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

We call (5) the *Best Reply Condition* for monomorphic distributions.

Since v^{BR} is independent of the length T of trust-building periods and $v(c_T; p_T) = v(c_T, c_T) = (1-\delta^{2T})d + \delta^{2T}c$ is a decreasing function of T, there is a lower bound to T above which (5) is satisfied. To compute the lower bound explicitly, for any T, define $\underline{\delta}(T)$ as the solution to $v(c_T; p_T) = v^{BR}$, or

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

Then (5) is satisfied if and only if $\delta \geqq \underline{\delta}(T)$. It is easy to see that

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

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

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

Then the *Best Reply Condition* (5) for monomorphic distributions is satisfied if and only if $T \geqq \underline{\tau}(\delta)$. The above argument is summarized as follows.

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

Two remarks are in order. First, $\underline{\tau}$ is a decreasing function of δ, since $\underline{\delta}$ is decreasing in T. Second, in the ordinary infinitely repeated Prisoner's Dilemma, the lower bound to the discount factor (as δ^2) that sustains the trigger-strategy equilibrium is $\sqrt{\frac{g-c}{g-d}} = \underline{\delta}(\infty)$. This means that cooperation in VSRPD requires more patience.

Next we investigate when a Nash equilibrium p_T is neutrally stable, i.e., c_T-strategy is a NSS. If a mutant s' can invade a distribution p, for any $\bar{\epsilon} > 0$, there exists $\epsilon \in (0, \bar{\epsilon})$ such that

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

By letting $\epsilon \to 0$, it must be an alternative best reply to p.

There are only two kinds of strategies that are possibly alternative best replies to p_T. The obvious ones are those that differ from c_T-strategy off the play path. These will give the same payoff as c_T-strategy and therefore cannot invade p_T. The other kind is the strategies that play D at some point when the partner is in the cooperation periods. When $T > \underline{\tau}(\delta)$, however, such strategies are not alternative best replies. Therefore c_T-strategy is a NSS for this case.

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

For any T, let $p_T^{T+1}(\alpha) = \alpha p_T + (1-\alpha)p_{T+1}$ be a two-strategy distribution of c_T and c_{T+1}. As we explained in the Introduction, the average payoff of c_T strategy, $v(c_T; p_T^{T+1}(\alpha))$, is strictly increasing and concave in α for any T, while the average payoff of c_{T+1}-strategy, $v(c_{T+1}; p_T^{T+1}(\alpha))$, is strictly increasing and convex when $T \leqq \underline{\tau}(\delta)$.

Lemma 3. *For any $\delta \in (\underline{\delta}(\infty), 1)$ and any $T = 0, 1, 2 \ldots$, $v(c_T; p_T^{T+1}(\alpha))$ is a strictly increasing and concave function of α.*

Proof: By differentiation. See Appendix.

Lemma 4. *For any $\delta \in (\underline{\delta}(\infty), 1)$ and any $T = 0, 1, 2 \ldots$ such that $T \leqq \underline{\tau}(\delta)$, $v(c_{T+1}; p_T^{T+1}(\alpha))$ is a strictly increasing and convex function of α.*

Proof: By differentiation. See Appendix.

By the definition of $\underline{\tau}(\delta)$, the average payoff of $c_{\underline{\tau}(\delta)}$ and $c_{\underline{\tau}(\delta)+1}$-strategy coincide at $\alpha = 1$. Thus, the convexity/concavity of the average payoff functions implies that $c_{\underline{\tau}(\delta)+1}$-strategy

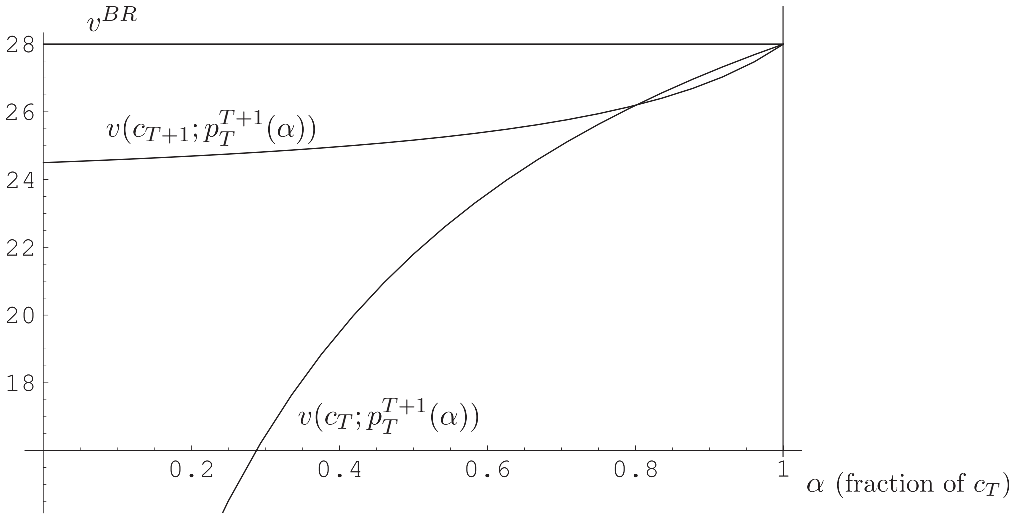

Figure 2: Average payoff functions of c_T-strategy and c_{T+1}-strategy when $T = \underline{\tau}(\delta)$ $(g = 60, c = 32, d = 0, \ell = -10, \delta = \frac{\sqrt{7/2}}{2}$, and $T = \underline{\tau}(\delta) = 1.)$

cannot invade $p_{\underline{\tau}(\delta)}$ if and only if the slope of $v(c_{\underline{\tau}(\delta)}; p_{\underline{\tau}(\delta)}^{\underline{\tau}(\delta)+1}(\alpha))$ is strictly smaller than the slope of $v(c_{\underline{\tau}(\delta)+1}; p_{\underline{\tau}(\delta)}^{\underline{\tau}(\delta)+1}(\alpha))$ at $\alpha = 1$. (See Figure 2.) This relationship of the slopes is warranted if $\underline{\tau}(\delta)$ is not too large, because as T becomes larger, the merit of starting cooperation one period earlier becomes smaller.

Lemma 5. *Take any* $\delta \in (\underline{\delta}(\infty), 1)$. *Let* $T = \underline{\tau}(\delta)$. *Then*

$$\frac{\partial v(c_T; p_T^{T+1}(\alpha))}{\partial \alpha}\bigg|_{\alpha=1} < \frac{\partial v(c_{T+1}; p_T^{T+1}(\alpha))}{\partial \alpha}\bigg|_{\alpha=1} \iff \left\{1 - \delta^{2(T+1)}\right\}(g - \ell) < c - d. \tag{6}$$

Proof: By computation. See Appendix.

Define $\hat{\tau}(\delta)$ implicitly as the solution to

$$\left\{1 - \delta^{2(T+1)}\right\}(g - \ell) = c - d. \tag{7}$$

Then $c_{\underline{\tau}(\delta)+1}$-strategy cannot invade $p_{\underline{\tau}(\delta)}$ if and only if $\underline{\tau}(\delta) < \hat{\tau}(\delta)$. To interpret (6), notice that $L(c_T, c_T) = 1 + \delta^2 + \cdots$ and $L(c_{T+1}, c_T) = 1 + \delta^2 + \cdots + \delta^{2T}$, so that $1 - \delta^{2(T+1)} = L(c_{T+1}, c_T)/L(c_T, c_T)$. Hence the condition (6) is equivalent to

$$(g - \ell)L(c_{T+1}, c_T) < (c - d)L(c_T, c_T) \tag{8}$$

at $T = \underline{\tau}(\delta)$. The RHS of (8) can be interpreted as the relative merit of c_T-strategy against c_{T+1}-strategy (to start cooperating one period early when meeting itself) and the LHS is the relative merit of c_{T+1}-strategy when meeting c_T-strategy.

As δ increases, T must increase to keep the equality (7). Thus $\hat{\tau}$ is a monotone increasing function of δ. Recall that $\underline{\tau}$ is a monotone decreasing function of δ. It is easy to show (see

Figure 3 in subsection 3.2) that there is a unique $\delta^* \in (\underline{\delta}(\infty), 1)$ such that

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

This δ^* is the critical survival rate such that (6) is satisfied at $T = \underline{\tau}(\delta)$ if and only if $\delta > \delta^*$. In summary, most of the monomorphic Nash equilibrium strategies are NSS except at boundary values when $\delta < \delta^*$.

Proposition 2. *(a) For any $\delta \in (\delta^*, 1)$, c_T-strategy is a NSS if and only if $T \geqq \underline{\tau}(\delta)$.*

(b) For any $\delta \in (\underline{\delta}(\infty), \delta^*]$, c_T-strategy is a NSS if and only if $T > \underline{\tau}(\delta)$.

It is possible to select among the monomorphic NSS's by cheap talk, under a slightly stronger definition of stability. The idea is that entrants can use a neologism at the beginning of a match to distinguish themselves from incumbents, imitate incumbents if the partner was an incumbent, and shorten trust-building periods if the partner was an entrant. If we require that entrants must be self-sustaining, i.e., the post-entry distribution must satisfy the *Best Reply Condition*, then the most efficient NSS is the unique strategy that is robust against self-sustaining entrants.

3.2 Bimorphic NSD

The nonlinearity of average payoff functions indicates that they may intersect when both strategies are present in the population, showing the potential for a *bimorphic* NSD consisting of two trust-building strategies.

Most literature on voluntarily separable repeated games has concentrated on monomorphic equilibria so that no voluntary break-up occurs, except for sorting out inherent defectors under incomplete information case. (See Section 5.) In this subsection we go beyond monomorphic equilibria and show the existence of bimorphic NSD. Since our model is of complete information and with homogeneous players, the following analysis can be interpreted as an evolutionary foundation of incomplete information models of diverse types of behaviors.

We focus on bimorphic distributions of the form $p_T^{T+1}(\alpha) := \alpha p_T + (1-\alpha)p_{T+1}$ and $T < \underline{\tau}(\delta)$. In order to compare efficiency with monomorphic NSDs, it is sufficient to investigate whether c_T-strategy with $T < \underline{\tau}(\delta)$ can be played by a positive measure of players. Against a c_T-strategy, c_{T+k}-strategies with $k \geqq 1$ behave the same way and, among those, c_{T+1}-strategy earns the highest payoff when meeting itself. Therefore for efficiency analysis (which is done in the next subsection 3.3), this class of bimorphic distributions are sufficient to consider.

For a bimorphic distribution to be a NSD, it needs to satisfy the following three conditions.

- All strategies in the support must earn the same average payoff.

- If the share of an incumbent strategy increases a little, its average payoff should be worse than the other strategy's and vice versa.

  Then the strategy distribution cannot be invaded by strategies that have the same play path as that of the incumbents.

- No strategy which differ in one step (on the play path) from some incumbent strategy can invade the distribution for sufficiently small ϵ.

The first two conditions can be jointly formulated as follows:

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

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

To deter invasion of mutants with one-step different strategy on the play path, we divide the mutants into two classes: those that play differently in the Prisoner's Dilemma and those that choose a different continuation decision.

First, consider mutants who play differently in Prisoner's Dilemma on the play path. At $t \leqq T$, the incumbents will play D. If a mutant plays C, it receives ℓ and goes back to the matching pool, while any incumbent receives $d(> \ell)$ and has less periods to build trust. Thus any incumbent has strictly larger continuation payoff than the mutant. At $T+1$, any action profile can occur on the play path, so we do not have to consider mutants. At $t \geqq T+2$, the incumbents will play C. If a mutant plays D and follows c_{T+k}-strategy ($k = 0, 1$) afterwards, its continuation payoff is $g + (L-1)v(c_{T+k}; p_T^{T+1}(\alpha_T^{T+1}))$, while c_{T+k} itself has the continuation payoff of

$$\frac{c}{1-\delta^2} + (L - \frac{1}{1-\delta^2})v(c_{T+k}; p_T^{T+1}(\alpha_T^{T+1})).$$

Combining with *Stable Payoff Equalization*, we can warrant that this type of mutants do strictly worse than the incumbents if

$$\begin{aligned} & g + (L-1)v(c_{T+k}; p_T^{T+1}(\alpha_T^{T+1})) < \frac{c}{1-\delta^2} + (L - \frac{1}{1-\delta^2})v(c_{T+k}; p_T^{T+1}(\alpha_T^{T+1})) \quad \forall k = 0, 1, \\ \iff \quad & v(c_T; p_T^{T+1}(\alpha_T^{T+1})) = v(c_{T+1}; p_T^{T+1}(\alpha_T^{T+1})) < v^{BR}. \end{aligned} \tag{10}$$

Note that if incumbents have strictly higher average payoff than a mutant, even if mutants enter with a positive measure, for sufficiently small measure, they cannot earn higher average payoff than any of the incumbents in the post-entry distribution.

Second, consider mutants who choose a different continuation decision on the play path. When a symmetric action profile is observed, incumbents would choose k. If a mutant ends the partnership, it goes back to the matching pool, and thus the continuation payoff is the payoff starting from a null history, while the incumbents have at least one period less to build trust or they are already in the cooperation periods, thus their continuation payoff is strictly larger.

When an asymmetric action profile is observed, incumbents would choose e. Even if a mutant wants to keep the partnership, the outcome cannot be changed in a match with an incumbent. In a match with another mutant, they do not observe an asymmetric action profile, thus their play path do not differ from the incumbents'.

In sum, a bimorphic NSD exists if there exists $\alpha_T^{T+1} \in (0,1)$ and its neighborhood such that (9) is satisfied and at that α_T^{T+1}, the average payoff is less than v^{BR}. For δ such that $\underline{\tau}(\delta) < \hat{\tau}(\delta)$ and T sufficiently close to but less than $\underline{\tau}(\delta)$, such α_T^{T+1} exists.

The idea of the proof is essentially that as T decreases slightly below $\underline{\tau}(\delta)$, the situation changes from Figure 2 to Figure 1b. Figure 2 shows that at $T = \underline{\tau}(\delta) < \hat{\tau}(\delta)$, the average payoff functions of c_T and c_{T+1} intersect at $\alpha = 1$ and at some $\alpha \in (0,1)$ thanks to the convexity/concavity. As T decreases from $\underline{\tau}(\delta)$, c_{T+1}-strategy has higher average payoff than c_T at $\alpha = 1$. However, since $T < \hat{\tau}(\delta)$ is warranted under $\delta > \delta^*$, the slope of $v(c_{T+1}; p_T^{T+1}(1))$ is steeper than the slope of $v(c_T; p_T^{T+1}(1))$, and hence by continuity of the average payoff functions with respect to T, there are two intersections in $(0,1)$, and the larger intersection satisfies *Stable Payoff Equalization* condition (9), as depicted in Figure 1b.

The strict *Best Reply Condition* (10) holds for any α which satisfies (9). If one-step deviation *after* $T+1$ is better than following the incumbents, then it is better to deviate *at* $T+1$. However, such strategy earns exactly the same payoff as that of c_{T+1}, which is also the same as that of c_T under the payoff equalization. Therefore no one-step deviation during cooperation periods of c_T or c_{T+1} earns higher payoff.

Proposition 3. *For any $\delta > \delta^*$, there exists $\tau_2(\delta) < \underline{\tau}(\delta)$ such that for any $T \in (\tau_2(\delta), \underline{\tau}(\delta))$[9] , there exists a bimorphic NSD with the support $\{c_T, c_{T+1}\}$.*

Proof: See Appendix.

Figure 3 illustrates the regions of (T, δ) where a monomorphic or a bimorphic NSD exist. To warrant an integer T, we need to restrict the payoff parameters G so that $(\tau_2(\delta), \underline{\tau}(\delta))$ contains an integer. Figure 1b is based on a numerical example[10] of such G that warrants a bimorphic NSD with the support $\{c_0, c_1\}$, even though c_0-strategy is never a NSS.

3.3 Higher Efficiency of Bimorphic NSD

For a given $\delta > \delta^*$, the shortest trust-building periods in the support of a bimorphic NSD is at least one period less than any of monomorphic NSS. Let the shortest trust-building periods

[9] As T decreases further, the intersection becomes unique. The unique intersection does not satisfy (9). Thus the sufficient range of T is an open interval.

[10] Namely, $g = 60, c = 55, d = 0, \ell = -1$, and $\delta = 0.95$.

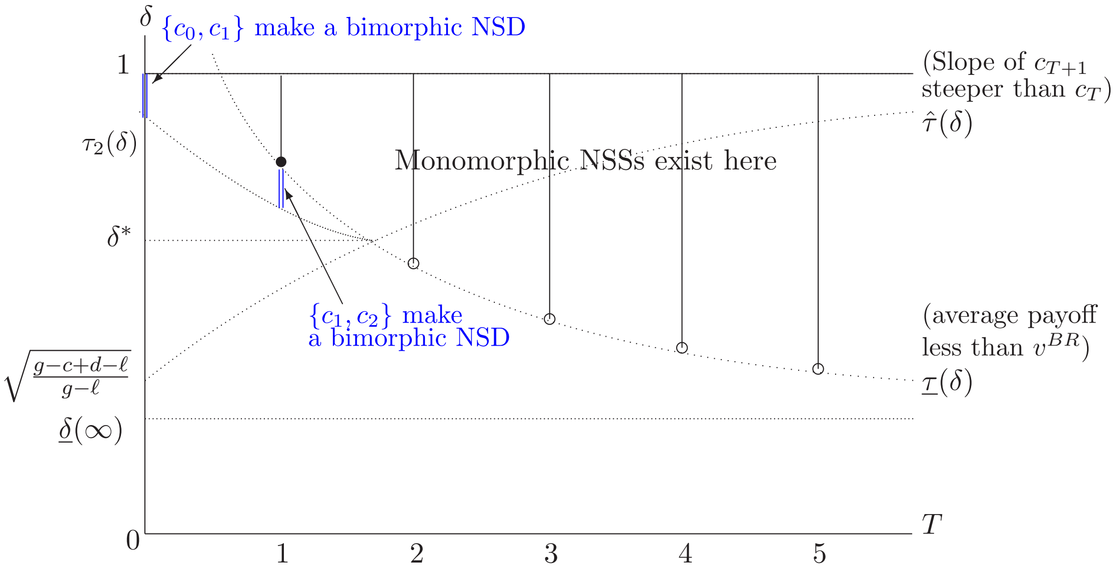

Figure 3: Parametric summary of monomorphic NSS and bimorphic NSD

of NSS be $T + 1$ and consider a bimorphic NSD with the support $\{c_T, c_{T+1}\}$. Let α_T^{T+1} be the fraction of c_T-strategy of the bimorphic NSD. The average payoff of c_{T+1} strategy as a NSS is

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

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

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

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

The intuition is as follows. Diverse strategies in the society make it valuable to maintain a relationship with the same-type partner, and if it is an equilibrium, the earliest cooperators (the "victims") must have enough share in the population to help each other. Thus a significant fraction of players can start cooperation early, which is more efficient than a monomorphic distribution under which all players must build trust initially.

Eeckhout (2006) shows that there exists a correlated strategy profile which Pareto dominates the best symmetric equilibrium, in a similar model to ours. His idea was to introduce a correlation device, such as skin color, to selectively start cooperation after a random match. By contrast, we have shown that correlation is not necessary to improve efficiency.

3.4 Polymorphic NSD

We can extend the analysis of the bimorphic NSDs to general *polymorphic* NSDs with more than two trust-building strategies in the support. However, finite-support NSDs are quite complex to analyze, while infinite-support NSDs are simpler. Let us explain this first and then focus on infinite-support NSDs.

Let us consider a trimorphic distribution with the support $\{c_T, c_{T+1}, c_{T+2}\}$. There are two ways to parameterize a trimorphic distribution. One parameterization is

$$\alpha p_T + \beta p_{T+1} + (1-\alpha-\beta) p_{T+2}, \tag{11}$$

while another way is

$$\alpha p_T + (1-\alpha)\gamma p_{T+1} + (1-\alpha)(1-\gamma) p_{T+2}, \tag{12}$$

where γ is the *relative* share of c_{T+1} as compared to that of c_{T+2}, given the fraction α of c_T-strategy. The parameterization (12) is easier to use, since we can decompose the *Stable Payoff Equalization* condition of three strategies into pairs, as follows. Given α, compare the payoffs of c_{T+1} and c_{T+2}. If these are equated and stability similar to (9) is satisfied, we can compare the payoffs of c_T and c_{T+1} (since c_{T+2} earns the same payoff as that of c_{T+1} against c_T). The stability similar to (9) for the payoff-equalizing $(\alpha_T^{T+2}, \gamma_{T+1}^{T+2})$ is formulated as follows. (The *Best Reply Condition* is derived in the same way as bimorphic NSD.)

In a neighborhood of $(\alpha_T^{T+2}, \gamma_{T+1}^{T+2})$, for any α, γ,

$$\alpha \gtreqless \alpha_T^{T+2} \iff v(c_{T+1}; p_T^{T+2}(\alpha, \gamma_{T+1}^{T+2})) \gtreqless v(c_T; p_T^{T+2}(\alpha, \gamma_{T+1}^{T+2})) \tag{13}$$

$$\gamma \gtreqless \gamma_{T+1}^{T+2} \iff v(c_{T+2}; p_T^{T+2}(\alpha_T^{T+2}, \gamma)) \gtreqless v(c_{T+1}; p_T^{T+2}(\alpha_T^{T+2}, \gamma)), \tag{14}$$

where

$$p_T^{T+2}(\alpha, \gamma) = \alpha p_T + (1-\alpha)\gamma p_{T+1} + (1-\alpha)(1-\gamma) p_{T+2}.$$

Note that change in α does not affect the relative share of c_{T+1} and c_{T+2}-strategies, and thus we require (13). Change in the share of c_{T+k} ($k \in \{1, 2\}$) affects the relative share γ, and thus we require (14).

It is probably not impossible to find a sufficient condition on T to warrant the existence of such $(\alpha_T^{T+2}, \gamma_{T+1}^{T+2})$ for the trimorphic case, but it is also easy to see that as the number of strategies in the support increases, the existence problem worsens because of more inequality conditions to satisfy. Therefore we do not pursue it here. (However, in Figure 4 we give a numerical example of a trimorphic NSD.)

A notable feature of trimorphic distributions is that the average payoff of c_{T+1}-strategy under a trimorphic distribution with the support $\{c_T, c_{T+1}, c_{T+2}\}$ is strictly less than that of c_{T+1} under a bimorphic distribution with the support $\{c_T, c_{T+1}\}$, since there is an exploiter

c_{T+2}-strategy. (See Figure 4.) This implies that if equilibrium α_T^{T+2} exists, then it is larger than the bimorphic equilibrium share α_T^{T+1}. Thus, if exists, a trimorphic NSD is more efficient than the bimorphic NSD with the same shortest trust-building periods T.

It turns out that strategy distributions with infinitely many trust-building strategies $\{c_T, c_{T+1}, \ldots\}$ in the support are easier to analyze, since they require only one parameter to equalize the payoffs. We first prove that the distribution must be "geometric" to equalize the payoffs of all strategies in $\{c_T, c_{T+1}, \ldots\}$.

Lemma 6. *For any $T < \infty$, let p be a stationary strategy distribution with the support $\{c_T, c_{T+1}, \ldots\}$. If $v(c_T; p) = v(c_{T+k}; p)$ for all $k = 1, 2, \ldots$, then there exists $\alpha \in (0, 1)$ such that the fraction of c_{T+k}-strategy is of the form $\alpha(1-\alpha)^k$ for each $k = 0, 1, 2, \ldots$.*

Proof: See Appendix.

Denote the geometric distribution of $\{c_T, c_{T+1}, \ldots\}$ as $p_T^\infty(\alpha)$. We show that if $p_T^\infty(\alpha)$ is the stationary strategy distribution in the matching pool, for any α, the average payoff of c_T-strategy is greater than/equal to/less than that of c_{T+1}-strategy if and only if the average payoff of c_{T+k}-strategy is greater than/equal to/less than that of c_{T+k+1}-strategy, for any $k = 1, 2, \ldots$. The intuition is as follows. The behavioral outcomes for c_{T+1}-strategy after the second period is essentially the same as those for c_T-strategy from the first period, i.e., T periods of trust-building followed by permanent cooperation if the partner had the same strategy, while followed by dissolution if the partner had a longer trust-building strategy. Similarly, the behavioral outcomes for c_{T+2}-strategy after the second period is essentially the same as those for c_{T+1}-strategy from the first period.... Therefore, if c_T has higher/the same/lower average payoff than c_{T+1}-strategy does, so does c_{T+1} against c_{T+2} and so on.

Lemma 7. *For any $T < \infty$ and any $\alpha \in (0, 1)$, $v(c_T; p_T^\infty(\alpha)) \gtreqless v(c_{T+1}; p_T^\infty(\alpha))$ if and only if $v(c_{T+k}; p_T^\infty(\alpha)) \gtreqless v(c_{T+k+1}; p_T^\infty(\alpha))$ for all $k = 1, 2, \ldots$.*

Proof: See Appendix.

A sufficient condition for one-step deviant mutants to be deterred is again the strict *Best Reply Condition*, which is derived in the same way as the bimorphic case. Notice that for any period after T, playing D (after the history consisting only of (D, D)) is an on-path action. Hence the meaningful deviations are those that play D after the cooperation periods started (that is, play D and keep the partnership for first $T + k$ periods, play C at least once, and then play D if the partnership continued). Such one-shot deviation during the cooperation periods cannot invade the distribution if the continuation value at $T + k + 2$ satisfies

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

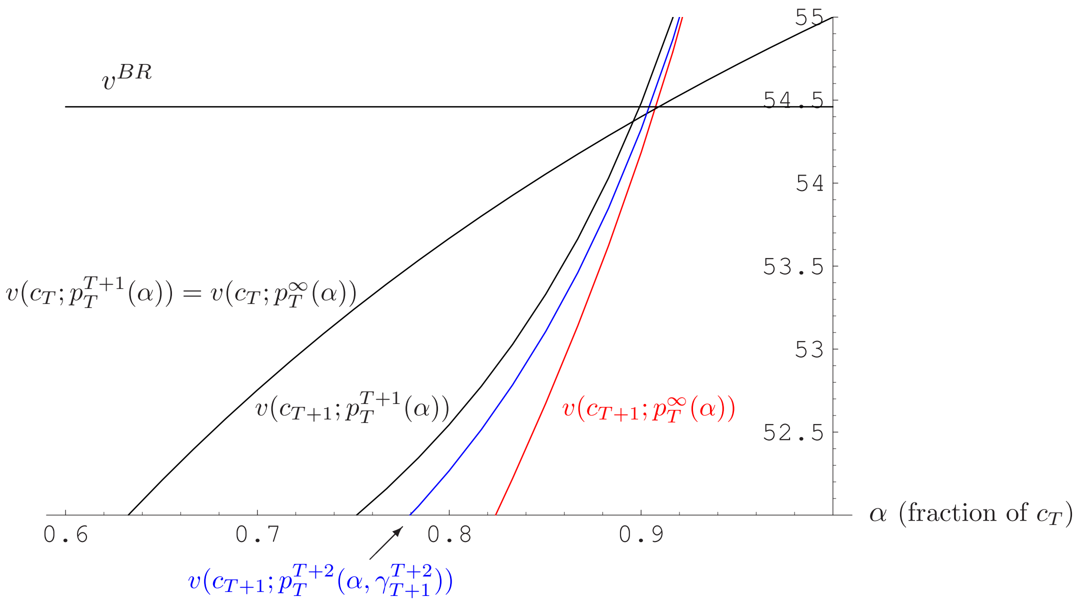

Figure 4: Bimorphic, Trimorphic, and Infinite-support NSD
($g = 60, c = 55, d = 0, \ell = -1, \delta = 0.95$, and $T = 0$.)

It is also straightforward to see that mutants who choose different continuation decision from the incumbents cannot invade the distribution.

Finally, mutants with the same play path as one of the c_{T+k}-strategies should be considered. However, stability under a geometric distribution is difficult to formulate, because change in the fraction of one strategy affects the relative share of many strategies. To simplify, we postulate that if the distribution changes from a geometric one, there will be evolutionary pressure to restore the distribution to another payoff-equalizing geometric distribution in a rather short time. Under this assumption, we show that there is a geometric distribution with $\alpha_T^{\infty} \in (0, 1)$ which is robust against small changes to another geometric distribution in the sense that there is a neighborhood U of α_T^{∞} such that for any $\alpha \in U$ and any $k = 1, 2, \ldots$,

$$\alpha \gtreqless \alpha_T^{\infty} \iff v(c_{T+k}; p_T^{\infty}(\alpha)) \gtreqless v(c_T; p_T^{\infty}(\alpha)). \tag{16}$$

The idea is that if mutants with the same play path as that of c_{T+k}-strategy enter, the distribution shifts to $p_T^{\infty}(\alpha)$ with $\alpha < \alpha_T^{\infty}$ and (16) warrants that they will do worse than the earliest cooperator c_T-strategy, and vice versa.

By a slight abuse of our terminology, we call an infinite-support distribution p a NSD if it satisfies (15) and (16). Note that if there is $\alpha_T^{\infty} \in (0, 1)$ that satisfies (16) for $k = 1$, then Lemma 7 warrants (16) for any $k = 1, 2, \ldots$.

Similar to the trimorphic case, the payoff of c_{T+1}-strategy is lower under an infinite-support distribution than that under a bimorphic distribution, since there are exploiters of c_{T+1}- strategy (namely, c_{T+k}-strategies with $k \geqq 2$). See Figure 4. Therefore the infinite-support NSD is more efficient than the bimorphic NSD with the same shortest trust-building period T.

Proposition 4. *For any $\delta > \delta^*$ there exists $\tau_\infty(\delta) < \underline{\tau}(\delta)$ such that for any $T \in (\tau_\infty(\delta), \underline{\tau}(\delta))$, there is a NSD of the form $p_T^\infty(\alpha_T^\infty)$ for some $\alpha_T^\infty \in (0,1)$. Moreover,*

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

i.e., the infinite-support NSD is more efficient than the bimorphic NSD with the same shortest trust-building periods.

Proof: See Appendix.

In summary, diverse strategies in the support improves the equilibrium average payoff. This is because the share of c_T-strategy (the earliest cooperators) must increase to restore the balance, when there are more exploiters in the distribution.

4 Extensions

4.1 Efficiency Wage and Matching Mechanism

Our model describes a society where players meet a stranger to play a Voluntarily Separable Repeated Prisoner's Dilemma. We analyzed how continuous cooperation becomes an equilibrium behavior when deviation from cooperation induces appropriate social sanctions.

Sanctions consist of two parts. First, a player's defection invokes the partner's severance decision, forcing him to start a new partnership with a stranger. Second, the payoff level he expects with a stranger is less than what he expects in a continued partnership with the current partner.

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

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

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

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

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

By the same logic as in subsection 3.1, the *Best Reply Condition* is $v^0(c_T; p_T, u) \leqq v^{BR}$. Clearly, if (5) is satisfied, this *Best Reply Condition* is also satisfied. Moreover, it can be satisfied even for c_0 for sufficiently large u, and cooperation without trust-building period becomes a self-sustaining state.[11]

As noted in Shapiro and Stiglitz (1984) and Okuno-Fujiwara (1989), unemployment works as a disciplinary device that deters moral hazard behavior. This observation suggests that the matching mechanism is an important element in creating cooperative partnerships.

There are two ways to consider the details of matching mechanisms. One way is to distinguish the reasons to be in the matching pool. In our setup, there are four reasons: new birth, death of the partner, separation due to the partner's deviation, and separation due to own deviation. In this paper we analyzed the case where no distinction can be made among these due to the lack of information. Higher efficiency is achieved if players can distinguish at least some reasons why the newly matched partner came into the matching pool. (See Okuno-Fujiwara et al., 2007.) Moreover, if players can find matches via their social network, before going to the random matching pool, further efficiency gain is expected since the deviation incentive is smaller within one's social network.

The other way is to find mechanisms that generate sufficient matching friction as assumed in the efficiency wage literature. Eeckhout (2006) shows that using personal characteristics (such as skin color) as a correlation device is a way to generate the sufficient matching friction in a monomorphic distribution.

4.2 Alternating-Action Equilibrium

If $2c < g + \ell$, then repeating (C, C) is not the most efficient outcome. Among pure strategy distributions, it is most efficient to alternate (C, D) and (D, C). By a similar logic to the monomorphic equilibrium, the following two-strategy distribution constitutes a NSD for sufficiently long trust-building periods.

Definition: For any $T = 1, 2, \ldots$, a_T-strategy is defined as follows.
$t = 1, 2, \ldots, T$: Play D and keep the partnership if and only if (D, D) is observed.
$t = T + 1$: Play C, keep the partnership regardless of the current observation, and move to *Alternating-action regime* if the partner played D. Move to *C-trigger regime* if the partner played C.
• *Alternating-action regime*: In periods $t = T + k$ ($k \geqq 2$) such that k is even, play D and keep the partnership if and only if the partner played C in the current period. When k is odd, play

[11] Carmichael and MacLeod (1997) have essentially the same idea by gift-giving instead of unemployment.

C and keep the partnership if and only if the partner played D.

• *C-trigger regime*: Play C and keep the partnership if and only if the partner played C in the current period.

Definition: For any $T = 1, 2, \ldots$, b_T-strategy is defined as follows.

$t = 1, 2, \ldots, T$: Play D and keep the partnership if and only if (D, D) is observed.

$t = T + 1$: Play D, keep the partnership regardless of the current observation, and move to *Alternating-action regime* if the partner played C. Move to *C-trigger regime* if the partner played D.

• *Alternating-action regime*: In periods $t = T + k$ ($k \geqq 2$) such that k is even, play C and keep the partnership if and only if the partner played D. When k is odd, play D and keep the partnership if and only if the partner played C.

• *C-trigger regime*: Play C and keep the partnership if and only if the partner played C in the current period.

If a_T met a_T, the play path is the same as c_T meeting c_T. If a_T met b_T, the play path after T periods of trust-building alternates action profiles (C, D) and (D, C). If b_T met b_T, the play path is the same as c_{T+1} meeting c_{T+1}. Therefore if a stationary distribution of a_T and b_T is stable, then a constant fraction of the population play the alternating action profiles after certain periods of trust building.

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

5 Conclusion and Related Literature

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

First, the trust-building periods in our equilibria serve as a mechanism for sanction against defection because they make the initial value of a new partnership small. In the literature, the gift exchange of Carmichael and MacLeod (1997) and the gradual cooperation in Datta (1996) and Kranton (1996a) have the same function. By contrast, the gradual cooperation under incomplete information (Ghosh and Ray, 1996, and Kranton, 1996a) is to sort types out and thus has a different meaning.[12]

[12] Repeated games with a quitting option (Watson, 2002, Blonski and Probst, 2001, and Furusawa and Kawakami, 2006) also display gradual cooperation to sort types.

Our model is based on more basic primitives than these previous works: the game is of complete information, the stage game is an ordinary prisoner's dilemma with two actions, and there is no gift exchange prior to the partnership. We show that it is still possible to construct a punishment mechanism. Furthermore, we consider evolution of behaviors within a society as a whole, rather than restricting attention to behaviors within a single partnership given a (monomorphic) strategy distribution in the society. We are also able to provide fuller characterizations of monomorphic trust-building strategy NSDs, such as identifying the condition (in terms of the survival rate and payoffs of stage game) for the existence of a NSD with a particular length of trust-building periods.

Eeckhout (2006) analyzes a very similar model to ours, except that his does not have a random death and he does not consider evolutionary stability. Since the most efficient monomorphic equilibrium does not generically attain the constrained optimal payoff v^{BR}, Eeckhout (2006) introduced a public randomization to improve the payoffs. By contrast, we noticed the nonlinearity of payoffs under asymmetric strategy distributions and thus did not have to resort to correlated strategies to improve the payoffs. In general, allowing correlated strategies requires the existence of a public randomization device, which is an extra assumption to the model.

Second, the existence and higher efficiency of polymorphic equilibria than monomorphic equilibria is a totally new result. The logic that early start of long-term cooperation is sustained because of possible exploitation in a future partnership is similar to the equilibrium of Rob and Yang (2005), written independently from our paper. In their model, there are three types of players; a bad type who always plays D, a good type who always plays C, and the rational type who tries to maximize their payoff. Existence of bad type players makes it valuable to (1) keep and cooperate with either good or rational type partners, and (2) to find out bad type partners as soon as possible. Thus, a rational player should cooperate from the beginning to be distinguished from the bad-type.

Our result is much starker than Rob and Yang's. Our model does not rely on heterogeneous "type" and incomplete information. Instead, bad (longer trust-building) strategy emerges endogenously as a polymorphic NSD. We also show that there are equilibria with a variety of (even infinitely many) heterogeneous strategies.

The higher efficiency of polymorphic equilibria is an interesting result. One might think that late cooperators create unstable partnerships and thus reduce social welfare. However, in equilibrium there should be enough early cooperators to help each other, and the benefit of partially early cooperation is greater than uniformly delayed cooperation.

6 Appendix: Proofs

Proof of Lemma 3: Let us rearrange $v(c_T; p_T^{T+1}(\alpha))$ to highlight the effect of α.

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

where

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

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

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

By differentiation,

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

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

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

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

Proof of Lemma 4: By the same logic as Lemma 3, let

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

Then

$$v(c_{T+1}; p_T^{T+1}(\alpha)) = v(c_{T+1}, c_{T+1}) + \mu_{T+1}(\alpha)\{v(c_{T+1}, c_T) - v(c_{T+1}, c_{T+1})\}. \tag{18}$$

Note that by computation

$$\begin{aligned}
& v(c_{T+1}, c_T) - v(c_T, c_T) \\
= & \frac{1}{1-\delta^{2(T+1)}}\Big[(1-\delta^{2T})d + \delta^{2T}(1-\delta^2)g - (1-\delta^{2(T+1)})(1-\delta^{2T})d - (1-\delta^{2(T+1)})\delta^{2T}c\Big] \\
= & \frac{\delta^{2T}}{1-\delta^{2(T+1)}}\Big[(1-\delta^2)(g-c) - \delta^2(1-\delta^{2T})(c-d)\Big] \gtreqqless 0,
\end{aligned} \tag{19}$$

if $T \leqq \underline{\tau}(\delta)$. (This fact will be useful in Proposition 3 as well.) Hence

$$\begin{aligned} & v(c_{T+1}, c_T) - v(c_{T+1}, c_{T+1}) \\ = \; & \{v(c_{T+1}, c_T) - v(c_T, c_T)\} + \{v(c_T, c_T) - v(c_{T+1}, c_{T+1})\} > 0, \end{aligned} \tag{20}$$

since c_T starts cooperation earlier than c_{T+1} (thus the second bracket is positive).

By differentiation,

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

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

Proof of Lemma 5: From (17) in the proof of Lemma 3 and (18) in the proof of Lemma 4,

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

As $\alpha \to 1$,

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

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

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

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

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

□

Proof of Proposition 3: We introduce a useful notation first. For any $T, T' \in \mathbb{N}$, define

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

Then for any $T, T' \in \mathbb{N}$ such that $T, T' \geqq 1$,

$$\begin{aligned}
\Gamma(c_T, c_{T'}) &= V(c_T, c_{T'}) - L(c_T, c_{T'})v^{BR} \\
&= d + \delta^2 V(c_{T-1}, c_{T'-1}) - \{1 + \delta^2 L(c_{T-1}, c_{T'-1})\}v^{BR} \\
&= d - v^{BR} + \delta^2 \Gamma(c_{T-1}, c_{T'-1}). \qquad (21)
\end{aligned}$$

Lemma 8. *For any $T, T' \in \mathbb{N}$, $\Gamma(c_T, c_T) = \Gamma(c_{T+1}, c_T)$.*

Proof of Lemma 8: We prove this by induction. The definition of v^{BR} is equivalent to

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

Hence we have that

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

That is,

$$\Gamma(c_0, c_0) = \Gamma(c_1, c_0).$$

Next suppose that $\Gamma(c_{T-1}, c_{T-1}) = \Gamma(c_T, c_{T-1})$ holds. From (21),

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

□

Lemma 9. *For any $\delta > \delta^*$, there exists $\tau_2(\delta) < \underline{\tau}(\delta)$ such that for any $T \in (\tau_2(\delta), \underline{\tau}(\delta))$, there are two solutions in $(0, 1)$ to*

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

Proof of Lemma 9: By the definition, at $T = \underline{\tau}(\delta)$,

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

Note also that for any T,

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

Recall that the average payoff of c_T is concave in α (Lemma 3), that of c_{T+1} is convex in α (Lemma 4), and $\delta > \delta^*$ warrants that the slope of the average payoff of c_{T+1} is steeper than that of c_T-strategy at $\alpha = 1$. Hence there are two intersections of the average payoff functions when $T = \underline{\tau}(\delta)$; one at $\alpha = 1$ and one within $(0, 1)$. (See Figure 2.)

From (19) we have that, for $T < \underline{\tau}(\delta)$,

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

but the slope of the average payoff of c_{T+1} is still steeper than that of c_T-strategy at $\alpha = 1$. Therefore, by the continuity of the average payoff functions with respect to T, for T sufficiently close to $\underline{\tau}(\delta)$, there are two intersections of the average payoff functions in $(0, 1)$. □

Let the larger solution be α_T^{T+1}. At α_T^{T+1}, the average value of c_{T+1} intersects with that of c_T from below. (See Figure 1b.) Hence there exists a neighborhood of α_T^{T+1} in which

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

To complete the proof of Proposition, we show that The *Best Reply Condition* is satisfied with strict inequality at α_T^{T+1}.

Let $\alpha_T^*(v^{BR})$ and $\alpha_{T+1}^*(v^{BR})$ be the fractions of c_T-strategy which solve $v(c_T; p_T^{T+1}(\alpha)) = v^{BR}$ and $v(c_{T+1}; p_T^{T+1}(\alpha)) = v^{BR}$ respectively. By the continuity of the average payoff functions with respect to T, it suffices to prove[13]

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

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

$$\begin{aligned}
& L(c_T; p_T^{T+1}(\alpha))\{v(c_T; p_T^{T+1}(\alpha)) - v^{BR}\} = 0 \\
\iff\ & V(c_T; p_T^{T+1}(\alpha)) - L(c_T; p_T^{T+1}(\alpha))v^{BR} = 0 \\
\iff\ & \alpha L(c_T, c_T)v(c_T, c_T) + (1-\alpha)L(c_T, c_{T+1})v(c_T, c_{T+1)} - L(c_T; p_T^{T+1}(\alpha))v^{BR} = 0 \\
\iff\ & \alpha\Gamma(c_T, c_T) + (1-\alpha)\Gamma(c_T, c_{T+1}) = 0.
\end{aligned}$$

Therefore, $\alpha_T^*(v^{BR})$ is the solution to

$$H_T^*(\alpha) := \alpha\Gamma(c_T, c_T) + (1-\alpha)\Gamma(c_T, c_{T+1}) = 0. \tag{22}$$

Similarly, $\alpha_{T+1}^*(v^{BR})$ is the solution to

$$H_{T+1}^*(\alpha) := \alpha\Gamma(c_{T+1}, c_T) + (1-\alpha)\Gamma(c_{T+1}, c_{T+1}) = 0. \tag{23}$$

By Lemma 8, $H_T^*(1) = \Gamma(c_T, c_T) = \Gamma(c_{T+1}, c_T) = H_{T+1}^*(1)$. Let us compare the slope of these linear functions:

$$\begin{aligned}
\frac{\partial H_T^*}{\partial \alpha} &= \Gamma(c_T, c_T) - \Gamma(c_T, c_{T+1}) \\
\frac{\partial H_{T+1}^*}{\partial \alpha} &= \Gamma(c_T, c_T) - \Gamma(c_{T+1}, c_{T+1}).
\end{aligned}$$

[13] $\alpha_{T+1}^*(v^{BR}) < \alpha_T^*(v^{BR})$ also holds when the smaller solution $\underline{\alpha}$ to $v(c_T; p_T^{T+1}(\alpha)) = v(c_{T+1}; p_T^{T+1}(\alpha))$ has the average payoff greater than v^{BR}. However, when $T = \underline{\tau}(\delta)$, $v(c_T; p_T^{T+1}(\underline{\alpha})) < v^{BR}$. Therefore for T sufficiently close to $\underline{\tau}(\delta)$, it cannot be that $v(c_T; p_T^{T+1}(\underline{\alpha})) > v^{BR}$.

In parameters,

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

Hence,

$$\begin{aligned}&\Big\{\Gamma(c_{T+1}, c_{T+1}) - \Gamma(c_T, c_{T+1})\Big\}(1-\delta^2)\\ =\ &\delta^{2T}(1-\delta^2)(d-\ell) + \delta^{2(T+1)}(c - v^{BR}) = \delta^{2T}(1-\delta^2)(d-\ell+g-c) > 0. \qquad (24)\end{aligned}$$

Therefore the slope of H^*_T is steeper than that of H^*_{T+1} for any α. This implies that $H^*_T(\alpha) < H^*_{T+1}(\alpha)$ for all $\alpha < 1$, and thus

$$\alpha^*_{T+1}(v^{BR}) < \alpha^*_T(v^{BR}).$$

This completes the proof of the proposition. □

Proof of Lemma 6: 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. (That is, the partnership has continued up to $t+1$-th period so that the possibility that the partner has a shorter trust-building than t is excluded.)

Let α_t be the conditional probability that the partner is the same strategy. The conditional probability is $1-\alpha_t$ that the partner has a longer trust-building period. The (non-averaged) continuation payoff of c_t-strategy at the beginning of $t+1$ is

$$V(c_t; p, t+1) = \alpha_t\{\frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2}V(c_t; p)\} + (1-\alpha_t)\{\ell + \delta V(c_t; p)\}. \qquad (25)$$

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

$$\begin{aligned}V(c_{t+1}; p, t+1) &= \alpha_t\{g + \delta V(c_{t+1}; p)\}\\ &\quad +(1-\alpha_t)\{d + \delta(1-\delta)V(c_{t+1}; p) + \delta^2 V(c_{t+1}; p, t+2)\}. \qquad (26)\end{aligned}$$

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 (26) becomes

$$\begin{aligned}V(c_{t+1}; p, t+1) &= \alpha_t\{g + \delta V(c_{t+1}; p)\}\\ &\quad +(1-\alpha_t)\{d + \delta(1-\delta)V(c_{t+1}; p) + \delta^2 V(c_t; p, t+1)\}\\ \Longleftrightarrow\ V(c_{t+1}; p, t+1) &= \frac{1}{1-(1-\alpha_t)\delta^2}\big[\alpha_t\{g + \delta V(c_{t+1}; p)\}\\ &\quad +(1-\alpha_t)\{d + \delta(1-\delta)V(c_{t+1}; p)\}\big]. \qquad (27)\end{aligned}$$

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

$$V(c_t;p) = V(c_{t+1};p). \tag{28}$$

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

$$V(c_t;p,t+1) = V(c_{t+1};p,t+1). \tag{29}$$

(29) implies that the RHS of (25) and (27) must be the same. Using (28) and letting $V^*(p) = V(c_t;p) = V(c_{t+1};p)$, α_t must satisfy

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

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

Proof of Lemma 7: For any $T, T' < \infty$, $\alpha \in [0,1]$ and $v \in \Re$, define

$$\begin{aligned}
\hat{\Gamma}(c_T, c_{T'}, v) &:= L(c_T, c_{T'})\{v(c_T, c_{T'}) - v\}; \\
\hat{\Gamma}(c_T, p_T^\infty(\alpha), v) &:= L(c_T; p_T^\infty(\alpha))\{v(c_T; p_T^\infty(\alpha)) - v\}.
\end{aligned}$$

Then, by (3),

$$\begin{aligned}
& L(c_T; p_T^\infty(\alpha))\{v(c_T; p_T^\infty(\alpha)) - v\} \\
= \; & \alpha L(c_T, c_T)v(c_T, c_T) + (1-\alpha)L(c_T, c_{T+1})v(c_T, c_{T+1}) - \{\alpha L(c_T, c_T) + (1-\alpha)L(c_T, c_{T+1})\}v \\
= \; & \alpha\hat{\Gamma}(c_T, c_T, v) + (1-\alpha)\hat{\Gamma}(c_T, c_{T+1}, v).
\end{aligned} \tag{30}$$

Similarly,

$$\begin{aligned}
& L(c_{T+1}; p_T^\infty(\alpha))\{v(c_{T+1}; p_T^\infty(\alpha)) - v\} \\
& = \alpha\hat{\Gamma}(c_{T+1}, c_T, v) + (1-\alpha)\{\alpha\hat{\Gamma}(c_{T+1}, c_{T+1}, v) + (1-\alpha)\hat{\Gamma}(c_{T+1}, c_{T+2}, v)\}; \qquad (31) \\
\text{and} \quad & L(c_{T+2}; p_T^\infty(\alpha))\{v(c_{T+1}; p_T^\infty(\alpha)) - v\} \\
& = \alpha\hat{\Gamma}(c_{T+2}, c_T, v) + (1-\alpha)[\alpha\hat{\Gamma}(c_{T+2}, c_{T+1}, v) \\
& \qquad + (1-\alpha)\{\alpha\hat{\Gamma}(c_{T+2}, c_{T+2}, v) + (1-\alpha)\hat{\Gamma}(c_{T+2}, c_{T+3}, v)\}], \qquad (32)
\end{aligned}$$

and so on. Note that by a generalization of (21), for any T, T' and $v \in \Re$,

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

Hence (31) becomes

$$
\begin{aligned}
& L(c_{T+1}; p_T^\infty(\alpha))\{v(c_{T+1}; p_T^\infty(\alpha)) - v\} \\
= \; & \alpha\hat{\Gamma}(c_{T+1}, c_T, v) + (1-\alpha)[d - v + \delta^2\{\alpha\hat{\Gamma}(c_T, c_T, v) + (1-\alpha)\hat{\Gamma}(c_T, c_{T+1}, v)\}] \\
= \; & \alpha\hat{\Gamma}(c_{T+1}, c_T, v) + (1-\alpha)[d - v + \delta^2\hat{\Gamma}(c_T, p_T^\infty, v)]. \qquad (33)
\end{aligned}
$$

Since c_{T+1} and c_{T+2} behave the same way against c_T, $\hat{\Gamma}(c_{T+2}, c_T, v) = \hat{\Gamma}(c_{T+1}, c_T, v)$. Thus (32) becomes

$$
\begin{aligned}
& L(c_{T+2}; p_T^\infty(\alpha))\{v(c_{T+1}; p_T^\infty(\alpha)) - v\} \\
= \; & \alpha\hat{\Gamma}(c_{T+1}, c_T, v) + (1-\alpha)\big[(d - v) + \delta^2\hat{\Gamma}(c_{T+1}, p_T^\infty(\alpha), v)\big]. \qquad (34)
\end{aligned}
$$

Subtracting (33) from (34), we get

$$
\hat{\Gamma}(c_{T+2}, p_T^\infty(\alpha), v) - \hat{\Gamma}(c_{T+1}, p_T^\infty(\alpha), v) = (1-\alpha)\delta^2\big\{\hat{\Gamma}(c_{T+1}, p_T^\infty(\alpha), v) - \hat{\Gamma}(c_T, p_T^\infty(\alpha), v)\big\}. \quad (35)
$$

Let $v = v(c_{T+1}; p_T^\infty(\alpha))$. Then $\hat{\Gamma}(c_{T+1}, p_T^\infty(\alpha), v) = 0$ so that (35) is equivalent to

$$
\begin{aligned}
& L(c_{T+2}; p_T^\infty(\alpha))\{v(c_{T+2}; p_T^\infty(\alpha)) - v(c_{T+1}; p_T^\infty(\alpha))\} \\
= \; & -(1-\alpha)\delta^2 L(c_T; p_T^\infty(\alpha))\{v(c_T; p_T^\infty(\alpha)) - v(c_{T+1}; p_T^\infty(\alpha))\}.
\end{aligned}
$$

Therefore, for any $\alpha \in (0,1)$,

$$
\begin{aligned}
& v(c_T; p_T^\infty(\alpha)) \gtreqless v(c_{T+1}; p_T^\infty(\alpha)) \\
\iff \; & v(c_{T+1}; p_T^\infty(\alpha)) \gtreqless v(c_{T+2}; p_T^\infty(\alpha)).
\end{aligned}
$$

This argument can be continued for any c_{T+k}-strategy. □

Proof of Proposition 4: We first prove the Stable Payoff Equalization and local stability of c_T and c_{T+1}-strategy, which also implies that all strategies in the support have the same average payoff and satisfy the stability by Lemma 7.

Lemma 10. *For any $\delta > \delta^*$, there exists $\underline{\tau}_\infty^E(\delta) < \underline{\tau}(\delta)$ such that for any $T \in (\underline{\tau}_\infty^E(\delta), \underline{\tau}(\delta))$, there exists $\alpha_T^\infty(\delta) \in (0,1)$ such that*

$$
\alpha \gtreqless \alpha_T^\infty \iff v(c_{T+1}; p_T^\infty(\alpha)) \gtreqless v(c_T; p_T^\infty(\alpha)). \qquad (36)
$$

Proof of Lemma 10: We use a similar logic to the one for the existence of bimorphic NSD, although the average payoff of c_{T+1}-strategy is no longer convex. First, notice that for any T, α,

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

The average payoff of c_{T+1} can be decomposed as follows.

$$\begin{aligned} v(c_{T+1};p_T^{\infty}(\alpha)) &= \mu_{10}(\alpha)v(c_{T+1},c_T)+\mu_{11}(\alpha)v(c_{T+1},c_{T+1})+\{1-\mu_{10}(\alpha)-\mu_{11}(\alpha)\}v(c_{T+1},c_{T+2}) \\ &= v(c_{T+1},c_{T+2})+\mu_{10}(\alpha)\{v(c_{T+1},c_T)-v(c_{T+1},c_{T+2})\} \\ &\quad +\mu_{11}(\alpha)\{v(c_{T+1},c_{T+1})-v(c_{T+1},c_{T+2})\}, \end{aligned}$$

where

$$\begin{aligned} \mu_{10}(\alpha) &:= \frac{\alpha L(c_{T+1},c_T)}{L(c_{T+1};p_T^{\infty}(\alpha))}; \\ \mu_{11}(\alpha) &:= \frac{(1-\alpha)\alpha L(c_{T+1},c_{T+1})}{L(c_{T+1};p_T^{\infty}(\alpha))}; \\ L(c_{T+1};p_T^{\infty}(\alpha)) &= \alpha L(c_{T+1},c_T)+(1-\alpha)\alpha L(c_{T+1},c_{T+1})+(1-\alpha)^2L(c_{T+1},c_{T+2}). \end{aligned}$$

By computation

$$\begin{aligned} \frac{\partial\mu_{10}}{\partial\alpha} &= \frac{L(c_{T+1},c_T)}{L(c_{T+1};p_T^{\infty}(\alpha))^2}\Big[\alpha^2\{L(c_{T+1},c_{T+1})-L(c_{T+1},c_{T+2})\}+L(c_{T+1},c_{T+2})\Big]>0, \\ \frac{\partial\mu_{11}}{\partial\alpha} &= \frac{L(c_{T+1},c_{T+1})}{L(c_{T+1};p_T^{\infty}(\alpha))^2}\Big[\alpha^2\{L(c_{T+1},c_{T+2})-L(c_{T+1},c_T)\}+(1-2\alpha)L(c_{T+1},c_{T+2})\Big]. \end{aligned}$$

Therefore, for any T,

$$\frac{\partial v(c_{T+1};p_T^{\infty}(\alpha))}{\partial\alpha}\Big|_{\alpha=1}=\frac{L(c_T,c_T)}{L(c_{T+1},c_T)}\{v(c_{T+1},c_T)-v(c_{T+1},c_{T+1})\}=\frac{\partial v(c_{T+1};p_T^{T+1}(\alpha))}{\partial\alpha}\Big|_{\alpha=1}.$$

Recall that by Lemma 5, for any $T<\hat{\tau}(\delta)$,

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

For any $\delta>\delta^*$, we have that $\underline{\tau}(\delta)<\hat{\tau}(\delta)$. Hence for any $T<\underline{\tau}(\delta)$,

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

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

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

and for any $T<\underline{\tau}(\delta)$,

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

Thus, by continuity of the average payoff with respect to T, for T sufficiently close to $\underline{\tau}(\delta)$, $v(c_{T+1};p_T^{\infty}(\alpha))$ crosses with $v(c_T;p_T^{\infty}(\alpha))$ from below, i.e., (36) holds. $\square$

Next, we show the *Best Reply condition* with strict inequality.

Lemma 11. *For any $\delta > \delta^*$ there exists $\tau_\infty(\delta) \in [\underline{\tau}_\infty^E(\delta), \underline{\tau}(\delta))$ such that for any $T \in (\tau_\infty(\delta), \underline{\tau}(\delta))$,*

$$v(c_T; p_T^\infty(\alpha_T^\infty(\delta))) < v^{BR}.$$

Proof of Lemma 11: Recall that $\alpha_T^*(v^{BR})$ is defined by

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

Define $\alpha_{T+1}^\infty(v^{BR})$ implicitly by $v(c_{T+1}; p_T^\infty(\alpha)) = v^{BR}$. Then it is sufficient to prove

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

Recall that $\alpha_T^*(v^{BR})$ is the (unique) solution to

$$H_T^*(\alpha) := \alpha\Gamma(c_T, c_T) + (1-\alpha)\Gamma(c_T, c_{T+1}) = 0. \tag{37}$$

By the same logic, $\alpha_{T+1}^\infty(v^{BR})$ is a solution to

$$F(\alpha) := \alpha\Gamma(c_{T+1}, c_T) + (1-\alpha)\alpha\Gamma(c_{T+1}, c_{T+1}) + (1-\alpha)^2\Gamma(c_{T+1}, c_{T+2}) = 0. \tag{38}$$

By computation,

$$\begin{aligned}
F(1) &= \Gamma(c_{T+1}, c_T) = \Gamma(c_T, c_T), \quad \text{from Lemma 8,} \\
F'(\alpha) &= \Gamma(c_T, c_T) + \Gamma(c_{T+1}, c_{T+1}) - 2\Gamma(c_{T+1}, c_{T+2}) - 2\alpha\{\Gamma(c_{T+1}, c_{T+1}) - \Gamma(c_{T+1}, c_{T+2})\}, \\
F'(0) &= \Gamma(c_T, c_T) + \Gamma(c_{T+1}, c_{T+1}) - 2\Gamma(c_{T+1}, c_{T+2}) > 0, \\
&\quad (\text{since } \; \Gamma(c_T, c_T) > \Gamma(c_{T+1}, c_{T+1}) > \Gamma(c_{T+1}, c_{T+2})) \\
F'(1) &= \Gamma(c_T, c_T) - \Gamma(c_{T+1}, c_{T+1}) > 0.
\end{aligned}$$

Hence F is a strictly increasing, concave function for $\alpha \in [0,1]$ and $F(\alpha) = 0$ has a unique solution in $[0,1]$. Note also that $H_T^*(0) = \Gamma(c_T, c_{T+1})$ and $H_T^*(1) = \Gamma(c_T, c_T) = F(1)$.

Since H_T^* is linear in α and F is concave, if the slope of H_T^* at $\alpha = 1$ is steeper than the slope of F at $\alpha = 1$, then for T close to but less than $\underline{\tau}(\delta)$, H_T^* intersects with the horizontal axis at a larger α than F does.[14] (See Figure 5.)

In fact, $F'(1) = \Gamma(c_T, c_T) - \Gamma(c_{T+1}, c_{T+1}) < \Gamma(c_T, c_T) - \Gamma(c_T, c_{T+1}) = H_T^{*'}(1)$ since $\Gamma(c_{T+1}, c_{T+1}) > \Gamma(c_T, c_{T+1})$ from (24). Therefore, for T sufficiently close to $\underline{\tau}(\delta)$ (and not less than $\underline{\tau}_\infty^E(\delta)$ so that $\alpha_T^\infty(\delta)$ exists), $\alpha_{T+1}^\infty(v^{BR}) < \alpha_T^*(v^{BR})$. □

Lemma 12. *The average payoff of c_{T+1}-strategy under $p_T^\infty(\alpha)$ is lower than the one under $p_T^{T+1}(\alpha)$ when $v(c_{T+1}; p_T^\infty(\alpha)) \leqq v^{BR}$.*

[14] For $\delta > \delta^*$, $H_T^*(0) = \Gamma(c_T, c_{T+1}) > \Gamma(c_{T+1}, c_{T+2}) = F(0)$. Thus H_T^* is not uniformly below F.

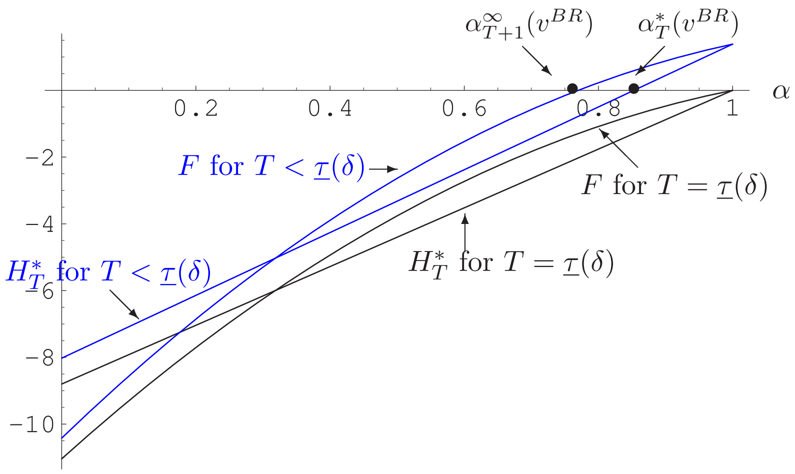

Figure 5: Proof of $\alpha^{\infty}_{T+1}(v^{BR}) < \alpha^{*}_{T}(v^{BR})$ for T sufficiently close to $\underline{\tau}(\delta)$

Proof of Lemma 12: For any $T \in N$ and any $(\alpha, v) \in (0,1) \times \Re$, define

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

Recall also

$$\begin{aligned}
\hat{\Gamma}(c_{T+1}, p^{\infty}_T(\alpha), v) &:= L(c_{T+1}; p^{\infty}_T(\alpha))\{v(c_{T+1}; p^{\infty}_T(\alpha)) - v\} \\
&= \alpha\Gamma(c_{T+1}, c_T) + (1-\alpha)\alpha\Gamma(c_{T+1}, c_{T+1}) + (1-\alpha)^2\Gamma(c_{T+1}, c_{T+2}) \\
&\quad +\Big\{\alpha L(c_{T+1}, c_T) + (1-\alpha)\alpha L(c_{T+1}, c_{T+1}) \\
&\qquad +(1-\alpha)^2 L(c_{T+1}, c_{T+2})\Big\}(v^{BR} - v). \qquad (39)
\end{aligned}$$

By computation, for any T,

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

Then for any $\alpha \in (0,1)$ and any $v \leqq v^{BR}$,

$$\begin{aligned}
\hat{\Gamma}(c_{T+1}, p^{T+1}_T(\alpha), v) - \hat{\Gamma}(c_{T+1}, p^{\infty}_T(\alpha), v) &= (1-\alpha)^2[\Gamma(c_{T+1}, c_{T+1}) - \Gamma(c_{T+1}, c_{T+2}) \\
&\quad +\{L(c_{T+1}, c_{T+1}) - L(c_{T+1}, c_{T+2})\}(v^{BR} - v)] > 0.
\end{aligned}$$

Now, consider

$$\begin{aligned}
& L(c_{T+1}; p^{T+1}_T(\alpha))[\{v(c_{T+1}; p^{T+1}_T(\alpha)) - v\} - \{v(c_{T+1}; p^{\infty}_T(\alpha)) - v\}] \\
= \;& \hat{\Gamma}(c_{T+1}, p^{T+1}_T(\alpha), v) - \hat{\Gamma}(c_{T+1}, p^{\infty}_T(\alpha), v) \\
& -\{L(c_{T+1}; p^{T+1}_T(\alpha) - L(c_{T+1}; p^{\infty}_T(\alpha))\}\{v(c_{T+1}; p^{\infty}_T(\alpha)) - v\}.
\end{aligned}$$

Note also $L(c_{T+1}; p_T^{T+1}(\alpha)) - L(c_{T+1}; p_T^{\infty}(\alpha)) = (1-\alpha)^2\{L(c_T, c_T) - L(c_{T+1}, c_{T+2})\} > 0$. Therefore,

$$v(c_{T+1}; p_T^{\infty}(\alpha)) \leqq v^{BR} \Rightarrow v(c_{T+1}; p_T^{T+1}(\alpha)) > v(c_{T+1}; p_T^{\infty}(\alpha)).$$

□

Lemma 12 implies that the infinite-support NSD is more efficient than the bimorphic NSD. This completes the proof of Proposition 4. □

References

Abreu, D. and R. Sethi. (2003): "Evolutionary Stability in a Reputational Model of Bargaining," *Games and Economic Behavior*, 44, 195–216.

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

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

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

Eeckhout, J. (2006): "Minorities and Endogenous Segregation," *Review of Economic Studies*, 73, 31-53.

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

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

Furusawa,T., and T. Kawakami (2006): "Gradual Cooperation in the Existence of Outside Options," forthcoming in *Journal of Economic Behavior & Organization.*

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

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

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

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

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

Okuno-Fujiwara, M., T. Fujiwara-Greve, and N. Suzuki (2007): "Social Norms and Voluntary Cooperation," (in Japanese) *The Economic Review (Keizai Kenkyu)*, 58, 110-121.

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

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

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

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