Tolerance and Behavioral Diversity*

Takako Fujiwara-Greve†
Dept. of Economics
Keio University

Masahiro Okuno-Fujiwara‡
Graduate School of Economics
University of Tokyo

March 11, 2019. Click here for the latest version.

Abstract: Globalization gives us a huge pool of potential transaction partners, but past behaviors of new partners are not easily observable, making it difficult to punish defectors. The literature of voluntarily separable repeated Prisoner's Dilemma advocates trust-building/gradual cooperation or co-existence of cooperators and defectors. Recent experiments reveal that subjects are more tolerant than the advocated equilibria. We show that tolerance is an equilibrium phenomenon, and moreover any action combination sequence is possible for arbitrary lengths of the "initial tolerant phase" of our equilibria. Behavioral diversity emerges under tolerance and provides incentives to cooperate as well. (95 words)

JEL code: C73.

Keywords: behavioral diversity, tolerance, cooperation, evolution, voluntary partnerships, repeated Prisoner's Dilemma.

*We thank Darrell Duffie, Hugo Sonnenschein, Bruno Strulovici, In-Koo Cho, Akihiko Matsui, Tadashi Sekiguchi, Takuo Sugaya, Chen-Lai Yang, and participants at Workshop "Dynamic Models in Economics" hosted by IMS (National University of Singapore) for helpful comments on various earlier drafts. Takako is grateful to Elizabeth Schwerer for checking the language and Stanford GSB for its hospitality where this study started. This work is supported by JSPS KAKENHI Grant Numbers (B) 24330064 and (B) 17H02501. The usual disclaimer applies.

†Corresponding author. Mailing address: Dept. of Economics, Keio University, 2-15-45 Mita, Minatoku, Tokyo 108-8345 JAPAN. Email: takakofg@econ.keio.ac.jp

‡Emeritus. Email: okuno.fujiwara@gmail.com

1 Introduction

In the age of globalization, markets no longer resemble repeated games in which players always interact with the same partners. Players can easily find and change partners via world-wide transportation and the internet. However, the past behavior of newly found partner(s) may not be perfectly observable, because there are always ways to erase one's past record such as changing locations or names.[1] The framework of voluntarily separable repeated Prisoner's Dilemma[2] (VSRPD) addresses this new opportunity of moral hazard.

When one can easily find a new partner who cannot know the opponent's past behavior, there is no Nash equilibrium in which all agents in the society act cooperatively towards a new partner. This is because, if all agents cooperate with strangers, a player who defects and changes partners every period will earn the highest one-shot payoff each time, i.e., starting every new match with cooperation is not a best response to itself.[3]

To sustain long-term cooperation, the literature proposes mainly two kinds of equilibria. One is trust-building/gradual-cooperation equilibria, in which partners do not fully cooperate initially but only after both players have persisted through low payoffs of mutual non-cooperation for a sufficient number of periods.[4] Defecting against a cooperative partner results in the slow process of trust-building with a new partner, which is the punishment. The other is the "fundamentally asymmetric" equilibrium (Fujiwara-Greve and Okuno-Fujiwara, 2012, Izquierdo et al., 2014, and Fujiwara-Greve et al., 2015) which consists of conditional cooperators (who always cooperate but stays with the partner if and only if mutual cooperation is observed) and myopic defectors (who always defect and leave immediately). In this equilibrium, the existence of myopic defectors gives incentives for conditional

[1] See for example, Datta (1996).

[2] See for example, Ghosh and Ray (1996), Kranton (1996), Carmichael and Macleod (1997), Eeckhout (2006), Fujiwara-Greve and Okuno-Fujiwara (2009), Rob and Yang (2010), and Mcadams (2011). Some of these use Prisoner's Dilemma with more than two actions. There are also dynamic analyses by Schumacher (2013) and Izquierdo et al. (2014), focusing on stationary or Markov strategies.

[3] All of the above models assume one-to-one matching. In a many-to-many matching model, Immorlica et al. (2014) showed that, if the one-shot deviation gain is small, the symmetric, fully cooperative equilibrium exists. The key is that, in their model, cooperators can accumulate multiple partners while myopic defectors cannot.

[4] See also Watson (2002) for an exit game which gives gradualism.

	previous period	(C,C)	(D,C)	(C,D)	(D,D)	new match
Okuno-Fujiwara et al. (2019)	Stay rate	0.999	0.699	0.427	0.458	
	C rate	0.963	0.510	0.587	0.3	0.3351
Lei et al. (estimated) (2018)	Stay rate	1	0.741	0.722	0.519	
	C rate	0.811	0.666	0.264	0.268	0.165
Lee (2018)	Stay rate	0.9983	0.906	0.636	0.6374	
	C rate	0.949	0.315	0.231	0.033	around 0.25
Honhon and Hyndman (2017)	(S,S) rate	0.998	0.804	0.655	0.36	
	(C,C) rate	0.9776	0.2063		0.016	0.65/0.37

Table 1: Experimental observations of VSRPD[5]

cooperators to stay with each other, which requires mutual cooperation.

Although trust-building/gradual cooperation and the co-existence of cooperators and defectors are both plausible in many contexts, we still observe different kinds of behavior in voluntary partnership situations. Table 1 shows summary statistics of some recent experiments of (slightly varied) VSRPD with no information flow across partnerships.[6] Three robust observations across these experiments are (i) that partnerships are not always terminated even if (C,C) is not played, (ii) that staying and cooperation rate are nearly 100% after mutual cooperation, and (iii) that both C and D are chosen after other action combinations as well as in new matches. (Table 1 should not be interpreted that Markov strategies are the most common behavior. See Okuno-Fujiwara et al., 2019.) Although the **efficient** behavior (ii) has been embedded in all existing equilibria, the **tolerant** behavior (i) and the **behavioral diversity** (iii) (in particular after mis-coordination) have never been addressed because all existing equilibria use immediate termination of a partnership as the main disciplining device (cf. Lei et al., 2018).

In this paper we construct new equilibria which are consistent with the observations (i) to (iii).

[5] We computed Lee's C rate from her Table 10 and Graph 2. (Pooled data of three treatments that allow voluntary separation.) The C rates of new matches of Honhon and Hyndman (2017) are individual C rates for the very first period of the dynamic game (0.65) and across all dynamic games (0.37) respectively.

[6] Okuno-Fujiwara et al. (2019) is most closely designed to implement the model of VSRPD using the **doubly-stochastic horizon** such that each partnership and the entire dynamic game are independently and randomly terminated every round. Lee (2018) and Honhon and Hyndman (2017) (TBA-U experiment) have only random termination of the dynamic game. Lei et al. (2018) is a finite horizon (40 rounds) game. Lei et al. (2018) only reports estimated action rates and Honhon and Hyndman (2017) do not report individual subjects' action rates except for the new matches.

The construction is by **tolerant** extensions of the conditional cooperator strategy and the myopic defector strategy: each player has some periods of initial **tolerant phase**, during which (s)he plays a pre-planned action sequence and does not voluntarily leave the match regardless of the partnership history, except when (C, C) is established, in which case (s)he becomes a conditional cooperator. After the tolerant phase is over, (s)he commits to either the conditional cooperator strategy or the myopic D-and-Leave strategy. This class of *tolerant strategies* describes human behaviors such that people want to "wait and see" how a new partnership develops before deciding to commit to a certain strategy. Since you are not fully committed to a certain behavior rule, you can also expect that the partner does not either and can tolerate her/his behavior for some time. As the strategies to commit, we simplify the analysis by focusing on the two simple but important strategies in the fundamentally asymmetric equilibrium of Fujiwara-Greve et al. (2015).

The existence of the fundamentally asymmetric equilibrium guarantees the existence of the tolerant equilibria (parameterized by the maximal length of the tolerant phase among the players). Therefore, voluntary relationships do not necessarily imply gradual cooperation nor a simple dichotomy of co-operators and defectors. Any action combination sequence can be observed for some periods, unless (C, C) is chosen. The model has a single homogeneous population facing an identical game, and thus we showed that heterogeneity among player characteristics is not needed for this behavioral diversity.

Tolerance and behavioral diversity support each other. On one hand, the tolerant phase allows initial behavioral diversity, because players stay after any partnership history with the hope to coordinate on mutual cooperation eventually or to exploit the partner at the end. On the other hand, behavioral diversity makes tolerance and long-term cooperation viable for two reasons. First, after any initial action history of a match, there is still a chance that the current partner cooperates in the future, because of the diverse strategy distribution in the matching pool from which the current partner comes. Second, going back to the random matching pool is risky because of the behavioral diversity as well.

The tolerant equilibria, with varied length of the initial tolerant phase, are all payoff-equivalent to one another, and the degenerate one is the fundamentally asymmetric equilibrium. Fujiwara-Greve et al. (2015) showed that the fundamentally asymmetric equilibrium is more efficient than any Nash equilibrium consisting of trust-building/gradual-cooperation strategies, under some payoff condition of the Prisoner's Dilemma. Hence the tolerant equilibria are also more efficient than any trust-building equilibrium under the same condition.

Since the game is an extensive-form game, it is desirable to investigate some refinements of Nash equilibria. We consider evolutionary stability, which is a strong refinement[7] of Nash equilibria by requiring robustness against positive measures of "mutants" who play different strategies from the equilibrium ones. Our model naturally embeds the underlying assumptions of evolutionary game models: it is a population game (Sandholm, 2010) and there is a death and birth process where natural selection as well as mutation/experimentation can take place. Moreover, evolutionary stability does not require that players know/have a consistent belief of the matching pool strategy distribution, which is also the case in experiments and the real world.

We show that tolerant equilibria are not less or more evolutionarily stable than trust-building equilibria. Specifically, tolerant equilibria are robust against "dispersed" mutants but not robust against coordinated pure-strategy mutants. By contrast, the (intolerant) trust-building equilibria are robust against any coordinated pure-strategy mutant, but not robust against some combinations of tolerant mutants, including the dispersed ones. We also note that Nash equilibria with only defecting players (tolerant or not) are less stable than both tolerant and trust-building equilibria.

In summary, tolerance and huge behavioral diversity, including cooperation with strangers, are equilibrium phenomena in a **homogeneous** population, even though players can defect and run away without information flow to future partners. This is an encouraging result for cooperative people under globalization. Moreover, there is no theoretical foundation that exclusion of heterogenous individuals from a population necessarily leads to a coordinated, cooperative society.

[7]For normal-form games, Evolutionarily Stable Strategy concept (Maynard Smith and Price, 1973 and Maynard Smith, 1982) is a stronger requirement than trembling-hand perfect equilibrium (because the ESS constitutes a proper equilibrium, see van Damme, 1987), and trembling-hand perfect equilibrium and sequential equilibrium are essentially the same (Selten, 1975).

	C	D
C	c, c	ℓ, g
D	g, ℓ	d, d

Table 2: Prisoner's Dilemma: $g > c > d > \ell$ and $2c \geqq g + \ell$.

The rest of the paper is organized as follows. In Section 2, we describe the model and the fundamentally asymmetric equilibrium of Fujiwara-Greve et al. (2015), which is the "basis" of the tolerant equilibria. In Section 3, we construct the tolerant equilibria and examine their properties. Section 4 concludes the paper.

2 Model and Preliminaries

2.1 Voluntarily Separable Repeated Prisoner's Dilemma

The model of *Voluntarily Separable Repeated Prisoner's Dilemma* (VSRPD) introduced by Fujiwara-Greve and Okuno-Fujiwara (2009) (henceforth Greve-Okuno) is a population game where randomly matched pairs of agents play the Prisoner's Dilemma (Table 2) repeatedly as long as both partners choose to stay. The time horizon is discrete, and the population is homogeneous and of size 1.

The time line (illustrated in Figure 1) and information structure are as follows. At the beginning of each period, each player either is matched with a partner from the previous period or enters a random matching process to find a new partner.[8] Newly matched players do not have information regarding one another's past actions. Matched partners play the symmetric Prisoner's Dilemma, their choices observable only to their current partners. After observing one another's actions in the Prisoner's Dilemma, the partners simultaneously choose between "Stay" and "Leave". A partnership dissolves if at least one partner chooses to leave. In addition, at the end of each period, players face an exogenous risk of exiting from the society, which we call "death", and survives to the next period only with probability $\delta \in (0, 1)$. If a player dies, a new player enters the society, keeping the population size constant. Newly born players and players who lose their partners either through death or by choice

[8]Following Greve-Okuno, we assume that the matching probability is 1. This makes cooperation most difficult, under the no-information-flow assumption below.

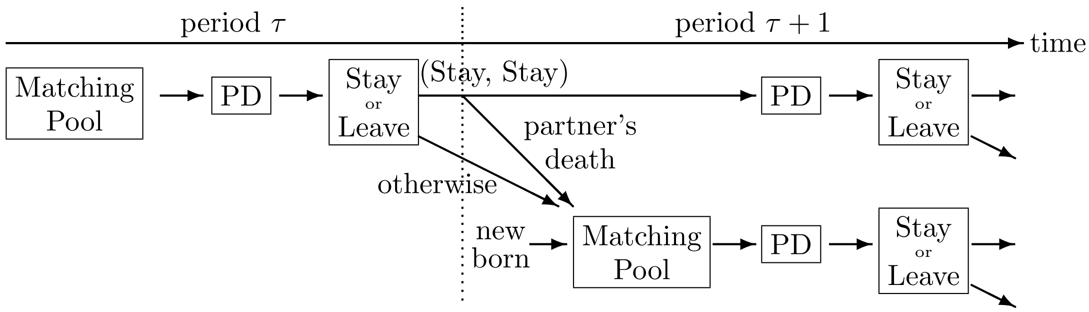

Figure 1: Timeline of the VSRPD

enter the matching pool in the next period. In sum, a partnership continues if and only if both partners live and choose to stay. In this case the partners play the Prisoner's Dilemma again in the next period, skipping the matching process. The game continues this way *ad infinitum*.

The one-shot payoffs in the Prisoner's Dilemma are in Table 2, where $g > c > d > \ell$ and $2c \geqq g + \ell$. The latter condition justifies our focus on (C, C) as the "focal" outcome of cooperative partnerships. Even if partners could alternate (C, D) and (D, C), it is not better than repeated mutual cooperation. Each individual player's game continues with probability δ, hence δ is the effective discount factor of a player. (However, even if both partners choose to stay in the partnership, the partnership continues only with probability δ^2.)

This model mimics a long-horizon, large market/society. Each player is so small that an individual player's strategy does not have an impact on the social distribution and her/his past behavior is difficult to verify to a randomly matched new partner. With the ease of finding a new partner, this model makes cooperation very difficult. In other words, if cooperative long-term partnerships are sustained in this model, they also exist in models with some information transmission and/or "unemployment" such that the probability of finding a new partner is less than 1 and waiting for a partner is costly.

2.2 Private strategies and match-independent strategies

In the VSRPD model, the largest class of pure strategies is the *private strategies*, which choose actions based on each player's private history from her/his birth until death. However, since the population is a continuum and there is no information flow across partnerships, the "contagious" strategies (Kandori, 1992 and Ellison, 1994) that change behavior towards a new partner based on one's private history with past partners cannot make an impact on a positive measure of the population and hence are irrelevant. Thus, although we define the private strategies for the completeness of the paper[9], we focus on equilibria consisting of *match-independent strategies* which base actions only on the (mutually observable) history within the same partnership.

For each player, let $\tau = 1, 2, 3\dots$ be her/his life periods starting at the "birth" into the matching pool (not the calendar time of the game nor the periods in a particular match). Let $\mathcal{H}_1 := \{\text{new}\}$ be the degenerate set of "private histories" that a newborn player has. For each $\tau = 2, 3, \dots$, let

$$\mathcal{H}_\tau = \{\text{new}\} \times [\{C, D\}^2 \times \{\text{Stay, Leave}\}^2 \times \{\text{new, continuing}\}]^{(\tau-1)}$$

be the set of private histories of a player, which records whether the match is a new one or a continuing one, and the action combinations within the experienced matches until the beginning of τ-th period of her/his life. (Recall that all players in the matching pool gets a partner every period. So each player observes a PD action combination and Stay/Leave choice combination every period with some partner. A player also learns whether the current match is a new one or a continuing one, but not the past action history of a new partner.)

Definition 1 A *(pure) private strategy* of a player is a sequence of cooperate/defect and stay/leave decision rules $s = (x_{s,\tau}, y_{s,\tau})_{\tau=1}^{\infty}$ such that for all $\tau = 1, 2, \dots$,

$x_{s,\tau} : \mathcal{H}_\tau \to \{C, D\}$;

$y_{s,\tau} : \mathcal{H}_\tau \times \{C, D\}^2 \to \{\text{Stay, Leave}\}$.

The set of pure private strategies is denoted by $\mathcal{S}$. In particular (and following the literature), we focus on the class of match-independent (private) strategies, defined below, to construct equilibria. Denote by $t = 1, 2, \dots$ the period within the same partnership. For each player, if a partnership dissolves, the next partnership starts at $t = 1$.

[9] This is also to clarify that a Nash equilibrium is defined as usual. We thank Michihiro Kandori for pointing out the need of this clarification.

Definition 2 For each $t = 2, 3 \ldots$, define the set of *partnership histories* as private histories of a single continuing match for t-periods:

$$H_t := \{\text{new}\} \times [\{C, D\}^2 \times \{(\text{Stay}, \ \text{Stay})\} \times \{\text{continuing}\}]^{t-1}$$

and the partnership history at $t = 1$ is the degenerate one; $H_1 = \{\text{new}\}$.

Definition 3 A *(pure) match-independent strategy* is a sequence of cooperate/defect and stay/leave decision rules $s = (x_{s,t}, y_{s,t})_{t=1}^{\infty}$ which only depend on the partnership periods $t = 1, 2, \ldots$ and partnership histories;

$x_{s,t} : H_t \to \{C, D\}$;

$y_{s,t} : H_t \times \{C, D\}^2 \to \{\text{Stay}, \ \text{Leave}\}$.

If a player uses a pure, match-independent strategy s, (s)he always chooses the same action $x_{s,1} \in \{C, D\}$ at the beginning of **any** new match, and follows s thereafter in **any** match. The set of pure, match-independent strategies is denoted by $\mathbf{S}$. Denote by $\mathcal{P}(\mathbf{S})$ the set of all probability distributions over $\mathbf{S}$. A strategy distribution $p \in \mathcal{P}(\mathbf{S})$ is interpreted that $p(s) \in [0, 1]$ of the players use the pure strategy s, for each s in the support of p, by the Law of Large Numbers (Sun, 2006). For notational simplicity, $s \in \mathbf{S}$ is also interpreted as the strategy distribution in $\mathcal{P}(\mathbf{S})$ that puts mass one on s.

From the evolutionary game perspective, we assume that each player is endowed with a pure strategy throughout her/his life, but the analysis does not change if each player randomizes among the equilibrium strategies at the birth and sticks to the realized pure-strategy for the rest of the life.

2.3 Average payoff function

Each strategy's long-run payoff should be measured from its birth into the matching pool until its random death. Hence we focus on the stability of stationary strategy distributions **in the matching pool**.[10] Stationarity is needed to explicitly compute the average long-run payoff of each strategy.

[10]The strategy distribution in the matching pool is not the same as the strategy distribution of the entire society, because the partnerships do not end simultaneously for all players. However, each stationary distribution in the matching pool induces a unique stationary distribution of all "states" of partnerships (classified by the partnership histories) in the society. See footnote 7 of Greve-Okuno.

For any $s, s' \in \mathbf{S}$, let $T(s, s')$ be the planned duration of the partnership between a player with strategy s and a player with strategy s', and let $U(s, s')$ be the total expected payoff for the s-player matched with an s'-player. Under the stationary distribution p in the matching pool, the probability of being matched with an s'-player is $p(s')$ every period, by the Law of Large Numbers of the dynamic random matching framework (Duffie et al., 2018[11]). Hence the *lifetime expected payoff* $V(s; p)$ of an s-player facing a stationary distribution p (with a countable support) in the matching pool is recursively formulated as

$$V(s;p) = \sum_{s' \in supp(p)} p(s') \Big[U(s,s') \ + [\delta(1-\delta)\{1+\delta^2+\cdots+\delta^{2(T\{s,s')-2\}}\} + \ \delta^{2\{T(s,s')-1\}} \cdot \delta] V(s;p) \Big].$$

To explain, the s-player loses the partner due to "death" before the $T(s, s')$-th period with probability $\delta(1-\delta)\{1+\delta^2+\cdots+\delta^{2\{T(s,s')-2)}\}$. With probability $\delta^{2\{T(s,s')-1\}} \cdot \delta$, the partnership successfully continues $T(s, s')$ periods and the s-player lives to the next period to go back to the matching pool. Denote the expected length of an (s, s')-pair by $L(s, s') = 1 + \delta^2 + \cdots + \delta^{2\{T(s,s')-1\}}$. Then,

$$\delta(1-\delta)\{1+\delta^2+\cdots+\delta^{2\{T(s,s')-2\}}\} + \ \delta^{2\{T(s,s')-1\}} \cdot \delta = 1-(1-\delta)L(s,s'),$$

where $(1 - \delta)L(s, s')$ is the probability that the s-player dies when the partnership could continue. Hence,

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

so that the *average lifetime expected payoff* of an s-player facing a stationary distribution p in the matching pool is

$$v(s;p) := (1-\delta)V(s;p) = \frac{\sum_{s' \in supp(p)} p(s') U(s,s')}{\sum_{s' \in supp(p)} p(s') L(s,s')}. \tag{1}$$

[11] Their "mutation" should be interpreted as changes of each player's "states" which is a combination of whether the player is a newborn or not and the strategy (s)he has.

Unless p is a symmetric distribution of a single pure-strategy, this average lifetime payoff is **not linear** in the share $p(s')$ of any strategy $s' \in supp(p)$. For later reference, denote the numerator by $U(s;p) := \sum_{s' \in supp(p)} p(s')U(s,s')$ and the denominator by $L(s;p) := \sum_{s' \in supp(p)} p(s')L(s,s')$. These are linear in the share of each strategy, and we have that

$$v(s;p) = \frac{U(s;p)}{L(s;p)}. \tag{2}$$

It is not easy to compare the average payoffs of different strategies, because they may differ not only in the numerator $U(s;p)$ but also in the denominator $L(s;p)$. The next lemma makes the comparison easy. For any stationary strategy distribution $p \in \mathcal{P}(\mathbf{S})$ in the matching pool and two pure strategies $s, \hat{s} \in \mathbf{S}$, define

$$\Delta U(\hat{s}, s; p) := U(\hat{s}; p) - U(s; p);$$

$$\Delta L(\hat{s}, s; p) := L(\hat{s}; p) - L(s; p).$$

These are linear in the shares of the strategies and are easy to compute.

Lemma 1 *For any $p \in \mathcal{P}(\mathbf{S})$, any $s \in \mathbf{S}$, and any $\hat{s} \in \mathbf{S}$,*

(i) if $\Delta L(\hat{s}, s; p) > 0$, then

$$v(s;p) > v(\hat{s};p) \iff v(s;p) > \frac{\Delta U(\hat{s}, s; p)}{\Delta L(\hat{s}, s; p)};$$

(ii) if $\Delta L(\hat{s}, s; p) < 0$, then

$$v(s;p) > v(\hat{s};p) \iff v(s;p) < \frac{-\Delta U(\hat{s}, s; p)}{-\Delta L(\hat{s}, s; p)}.$$

Proof. See Appendix.

An intuition for Lemma 1 (i) is that, when $\Delta L(\hat{s}, s; p) > 0$, the "less tolerant" s-strategy performs better if and only if the average extra payoff $\frac{\Delta U(\hat{s},s;p)}{\Delta L(\hat{s},s;p)}$ that the "more tolerant" $\hat{s}$-strategy gets by its tolerance is less than the payoff $v(s;p)$ that s gets by going back to the matching pool. (Since s and $\hat{s}$ are facing the same strategy distribution p, longer $L(\hat{s};p)$ means that $\hat{s}$ does not leave some partnerships that s leaves.) Lemma 1 (ii) is analogously interpreted.

Sometimes, the ratio $\frac{\Delta U(\hat{s},s;p)}{\Delta L(\hat{s},s;p)}$ has many terms in the numerator and the denominator. We have a lemma to simplify the computation of $\frac{\Delta U(\hat{s},s;p)}{\Delta L(\hat{s},s;p)}$ as well.

Lemma 2 *For any finite $J \in \mathbb{N}$, any $A_1, \ldots, A_J \in \mathbb{R}$, and any $B_1, \ldots, B_J \in \mathbb{R}_{++}$,*

$$\min_{j=1,\ldots,J}\{\frac{A_j}{B_j}\} \leqq \frac{\sum_{j=1}^{J} A_j}{\sum_{j=1}^{J} B_j} \leqq \max_{j=1,\ldots,J}\{\frac{A_j}{B_j}\}.$$

Proof. See Appendix.

2.4 Fundamentally asymmetric equilibrium

Before going to the construction of tolerant equilibria, we review the *fundamentally asymmetric equilibrium* of Fujiwara-Greve et al. (2015).

Definition 4 Let c_0*-strategy* be a match-independent strategy[12] as follows: in any period $t = 1, 2, \ldots$ of a partnership and after any partnership history, play C and, after that, stay if and only if (C, C) is observed in that period.[13]

Let d_0*-strategy* be a match-independent strategy as follows: in any period $t = 1, 2, \ldots$ and after any partnership history, play D and leave.[14]

The c_0-strategy is similar to the C-trigger strategy in the ordinary repeated Prisoner's Dilemma, but leaving the partnership is the punishment. The d_0-strategy is the most myopic strategy. The d_0-strategy constitutes a symmetric Nash equilibrium for any $\delta \in (0, 1)$[15], but the c_0-strategy does not. An interesting property of the VSRPD is that these two strategies can make a Nash equilibrium.

Definition 5 A stationary strategy distribution $p \in \mathcal{P}(\mathbf{S})$ in the matching pool is a *Nash equilibrium* if, for all $s \in supp(p)$ and all $s' \in \mathcal{S}$[16],

$$v(s;p) \geqq v(s';p). \tag{3}$$

[12] To be precise, this is a class of strategies, because we allow any off-path action plan in the information sets which are not reachable. The same caveat applies to other definitions of specific strategies.

[13] Formally, for any $t = 1, 2, \ldots$ and any $h \in H_t$, $x_{c_0,t}(h) = C$, $y_{c_0,t}(h, (C, C)) =$ Stay, and $y_{c_0,t}(h, (a, a')) =$ Leave for any $(a, a') \neq (C, C)$.

[14] Formally, for any $t = 1, 2, \ldots$ and any $h \in H_t$, $x_{d_0,t}(h) = D$, $y_{d_0,t}(h, (a, a')) =$ Leave for any $(a, a') \in \{C, D\}^2$.

[15] See Greve-Okuno Section 2.3, where it is called the $\tilde{d}$-strategy.

[16] Although we did not explicitly derive $v(s'; p)$ for each **private** strategy $s' \in \mathcal{S}$, it is possible to compute the continuation payoffs of all relevant one-step deviations, which are not restricted to be match-independent ones.

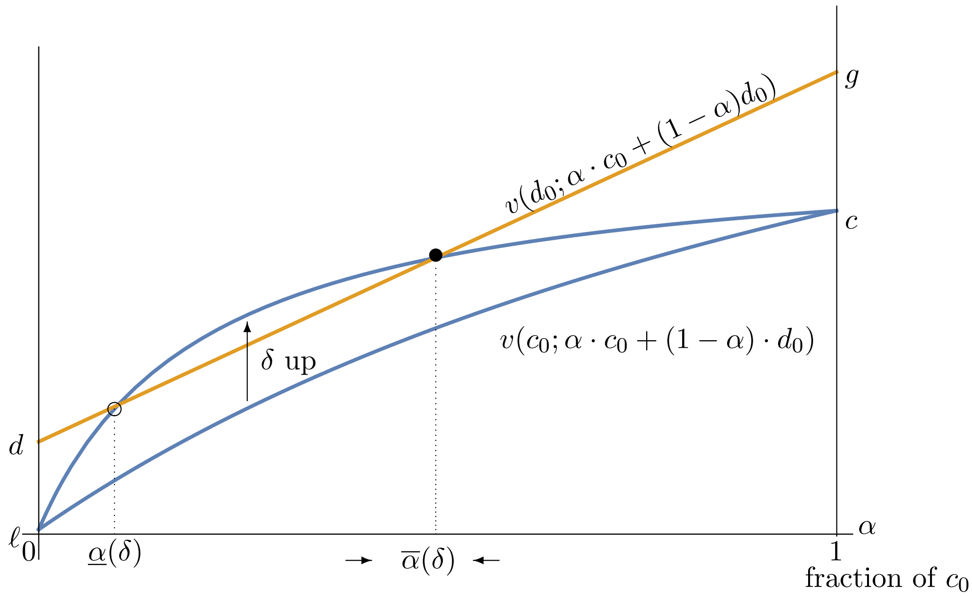

Figure 2: Locally stable c_0-d_0 equilibrium

Lemma 3 *(Fujiwara-Greve and Okuno-Fujiwara (2012), Fujiwara-Greve et al. (2015) and Izquierdo et al. (2014)) There exists* $\underline{\delta}_{c_0 d_0} \in (0,1)$ *such that* $\delta > \underline{\delta}_{c_0 d_0}$ *if and only if there is a unique*[17] $\overline{\alpha}(\delta) \in (0,1)$*, such that the bimorphic distribution* $p_{c_0 d_0}(\delta) = \overline{\alpha}(\delta) \cdot c_0 + \{1 - \overline{\alpha}(\delta)\} \cdot d_0$ *is a Nash equilibrium with the following "local stability" property: for each* $s \in \{c_0, d_0\}$*, there exists* $\bar{\epsilon} \in (0,1)$ *such that, for any* $\epsilon \in (0, \bar{\epsilon})$*,*

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

where $s' \neq s$ *is the other strategy in* $\{c_0, d_0\}$*.*

Figure 2 illustrates the intuition behind the payoff-equivalence of the c_0- and the d_0-strategy for sufficiently large δ, and the local stability. For notational convenience, for each $z \in \{c_0, d_0\}$ and any $\alpha \in [0,1]$, denote

$$U_0(z;\alpha) := U(z; \alpha \cdot c_0 + (1-\alpha)d_0);$$

$$L_0(z;\alpha) := L(z; \alpha \cdot c_0 + (1-\alpha)d_0);$$

$$v_0(z;\alpha) := \frac{U_0(z;\alpha)}{L_0(z;\alpha)} = v(z; \alpha \cdot c_0 + (1-\alpha)d_0).$$

[17] The other payoff-equivalent ratio of the c_0- and the d_0-strategy does not make the bimorphic distribution locally stable.

Explicitly, the average payoff of the c_0-strategy is

$$v_0(c_0;\alpha) = \frac{\alpha \cdot U(c_0,c_0) + (1-\alpha)U(c_0,d_0)}{\alpha \cdot L(c_0,c_0) + (1-\alpha)L(c_0,d_0)} = \frac{\alpha \cdot \frac{c}{1-\delta^2} + (1-\alpha)\ell}{\alpha \cdot \frac{1}{1-\delta^2} + (1-\alpha)}. \tag{4}$$

Hence it is monotone increasing and concave in α. By contrast, the average payoff of the d_0-strategy is linear in α:

$$v_0(d_0;\alpha) = \frac{\alpha \cdot U(d_0,c_0) + (1-\alpha)U(d_0,d_0)}{\alpha \cdot L(d_0,c_0) + (1-\alpha)L(d_0,d_0)} = \alpha \cdot g + (1-\alpha)d. \tag{5}$$

The concavity of $v_0(c_0;\alpha)$ is due to the voluntary nature of partnerships. As the survival rate δ increases, the average payoff of the c_0-strategy increases for any α, because c_0-pairs last longer. Hence, the average payoff function of the c_0-strategy becomes more concave as δ increases. For sufficiently high δ's, the average payoff functions of the two strategies have two intersections and the one with the larger share of the c_0-strategy is locally stable. Let us call this share $\overline{\alpha}(\delta)$ the *C-D ratio.*

Payoff equivalence of the c_0- and the d_0-strategy implies that they constitute a Nash equilibrium, by the following Lemma shown in Fujiwara-Greve et al. (2015). (This lemma will be useful later.)

Lemma 4 *(Lemma 5 of Fujiwara-Greve et al. (2015)) Fix an arbitrary* $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$ *and let* $v^M := v_0(c_0;\overline{\alpha}(\delta)) = v_0(d_0;\overline{\alpha}(\delta))$*. The common payoff* v^M *satisfies the Best Reply Condition in Greve-Okuno with the strict inequality:*

$$g + \delta\frac{v^M}{1-\delta} < \frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2} \cdot \frac{v^M}{1-\delta}. \tag{6}$$

The LHS of (6) is the continuation payoff of a one-step deviation from the c_0-strategy to play D when it is certain that the partner also has the c_0-strategy (i.e., in $t \geqq 2$ and after h with only (C,C) has been observed), and the RHS is the continuation payoff of the c_0-strategy, with random death of the partner included. It can be shown that other one-step deviations from either the c_0- or d_0-strategy after some partnership history are not beneficial, either. Therefore, for each $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, the bimorphic distribution $p_{c_0 d_0}(\delta) = \overline{\alpha}(\delta) \cdot c_0 + \{1 - \overline{\alpha}(\delta)\} \cdot d_0$ is a Nash equilibrium.

3 Tolerant Equilibria

3.1 Example: One-period tolerant equilibria

To give an intuition of our main result, in this subsection we explicitly construct a class of Nash equilibria in which some players do not end a partnership at the end of $t = 1$ regardless of its history (corresponding to the observation (i) in Introduction) and, if (C, C) is established, **all** players stay and play C again (corresponding to the observation (ii)). Intuitively, in the first period of a new match, players "take a look" at each other and not fully committed to behave cooperatively like the c_0-strategy or myopically like the d_0-strategy. While "taking a look", players can differ in the initial action, and that is tolerated.

Definition 6 For any (initial) action $X \in \{C, D\}$ and any strategy $z \in \{c_0, d_0\}$, the 1-*period tolerant strategy*, denoted X_z, is a match-independent strategy such that
if $t = 1$ (Tolerant phase): play X and stay for any observation;
if $t = 2$ (Commitment phase): play the c_0-strategy as the continuation strategy if (C, C) was observed in $t = 1$, and play the z-strategy otherwise.

Specifically, there are four 1-period tolerant strategies $C_{c_0}, C_{d_0}, D_{c_0}$ and D_{d_0}. Only the C_{d_0}-strategy has two possible continuation strategies in $t = 2$, depending on whether the $t = 1$ outcome was (C, C) or (C, D) (the first coordinate is the relevant player's action). The c_0- and the d_0-strategy can be also interpreted as degenerate 0-period tolerant strategies.

Fix an arbitrary $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$ and consider a class of strategy combinations of the form

$$\begin{aligned}\bar{p}_1 = &\overline{\alpha}(\delta)[\beta \cdot c_0 + (1-\beta)\{\overline{\alpha}(\delta) \cdot C_{c_0} + (1-\overline{\alpha}(\delta))C_{d_0}\}] \\ &+ (1-\overline{\alpha}(\delta))[\gamma \cdot d_0 + (1-\gamma)\{\overline{\alpha}(\delta) \cdot D_{c_0} + (1-\overline{\alpha}(\delta))D_{d_0}\}],\ \exists \beta, \gamma \in [0,1]. \qquad (7)\end{aligned}$$

Equivalently, we can write (7) as follows: all of the relative ratios between C-start strategies (c_0, C_{c_0}, C_{d_0}) and D-start strategies (d_0, D_{c_0}, D_{d_0}), between C_{c_0} and C_{d_0}, and between D_{c_0} and D_{d_0} are the C-D ratio, i.e., (with notational simplification such that $p(s, s', \ldots,) := p(\{s, s', \ldots\})$)

$$\bar{p}_1(c_0, C_{c_0}, C_{c_0}) = \overline{\alpha}(\delta), \quad \forall X \in \{C, D\},\ \bar{p}_1(X_{c_0}, X_{d_0}) > 0 \Rightarrow \frac{\bar{p}_1(X_{c_0})}{\bar{p}_1(X_{c_0}, X_{d_0})} = \overline{\alpha}(\delta). \qquad (8)$$

you \ partner	c_0	C_{c_0}	C_{d_0}	d_0	D_{c_0}	D_{d_0}
	C-start ($\overline{\alpha}(\delta)$)			D-start ($1-\overline{\alpha}(\delta)$)		
		$\overline{\alpha}(\delta) : 1-\overline{\alpha}(\delta)$			$\overline{\alpha}(\delta) \quad : 1-\overline{\alpha}(\delta)$	
c_0	$c, c, \ldots$	$c, c, \ldots$	$c, c, \ldots$	ℓ	ℓ	ℓ
C_{c_0}	$c, c, \ldots$	$c, c, \ldots$	$c, c, \ldots$	ℓ	$\ell, c, c, \ldots$	ℓ, ℓ
C_{d_0}	$c, c, \ldots$	$c, c, \ldots$	$c, c, \ldots$	ℓ	ℓ, g	ℓ, d
d_0	g	g	g	d	d	d
D_{c_0}	g	$g, c, c, \ldots$	g, ℓ	d	$d, c, c, \ldots$	d, ℓ
D_{d_0}	g	g, g	g, d	d	d, g	d, d

Table 3: Within-partnership payoff sequences in $\bar{p}_1$

There is freedom regarding whether $\bar{p}_1(X_{c_0}, X_{d_0}) > 0$ or not, for each $X \in \{C, D\}$, corresponding to whether $\beta, \gamma < 1$ or not. Therefore, this class includes the two-strategy equilibrium of $p_{c_0 d_0}(\delta)$ and distributions with three to six strategies in the support. Table 3 shows the payoff sequences of all possible partnerships and the relative ratio structure, which helps to understand the payoff equivalence of all constituent strategies in $\bar{p}_1$.

Proposition 1 *For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$ and any $\beta, \gamma \in [0, 1]$, $\bar{p}_1$ defined by (7) or (8) is a Nash equilibrium and is payoff-equivalent to $p_{c_0 d_0}(\delta)$.*

Proof of Proposition 1. Fix an arbitrary $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$ and arbitrary $\beta, \gamma \in [0, 1]$.

Step 1: $v(c_0; \bar{p}_1) = v(d_0; \bar{p}_1)$.

Proof of Step 1: From Table 3,

$$v(c_0; \bar{p}_1) = \frac{\bar{p}_1(c_0, C_{c_0}, C_{d_0})\frac{c}{1-\delta^2} + \bar{p}_1(d_0, D_{c_0}, D_{d_0})\ell}{\bar{p}_1(c_0, C_{c_0}, C_{d_0})\frac{1}{1-\delta^2} + \bar{p}_1(d_0, D_{c_0}, D_{d_0}) \cdot 1} = \frac{\overline{\alpha}(\delta) \cdot \frac{c}{1-\delta^2} + (1-\overline{\alpha}(\delta))\ell}{\overline{\alpha}(\delta) \cdot \frac{1}{1-\delta^2} + (1-\overline{\alpha}(\delta))} = v_0(c_0; \overline{\alpha}(\delta));$$
$$v(d_0; \bar{p}_1) = \frac{\bar{p}_1(c_0, C_{c_0}, C_{d_0}) \cdot g + \bar{p}_1(d_0, D_{c_0}, D_{d_0}) \cdot d}{\bar{p}_1(c_0, C_{c_0}, C_{d_0}) \cdot 1 + \bar{p}_1(d_0, D_{c_0}, D_{d_0}) \cdot 1} = \overline{\alpha}(\delta) \cdot g + (1-\overline{\alpha}(\delta))d = v_0(d_0; \overline{\alpha}(\delta)).$$

That is, the average payoff of the c_0-strategy (resp. d_0-strategy) under $\bar{p}_1$ is the same as the one under the fundamentally asymmetric equilibrium $p_{c_0 d_0}(\delta)$. By the definition of the C-D ratio in Lemma 3, $v(c_0; \bar{p}_1) = v_0(c_0; \overline{\alpha}(\delta)) = v_0(d_0; \overline{\alpha}(\delta)) = v(d_0; \bar{p}_1)$. □

Step 2: For each $s \in \{C_{c_0}, C_{d_0}\}$, $v(s; \bar{p}_1) = v(c_0; \bar{p}_1)$.

Proof of Step 2: From Table 3,

$$
\begin{aligned}
&U(C_{c_0};\bar{p}_1)\Big(=\sum_{s\in supp(\bar{p}_1)}\bar{p}_1(s)U(C_{c_0},s)\Big)\\
&=\bar{p}_1(c_0,C_{c_0},C_{d_0})\frac{c}{1-\delta^2}+\bar{p}_1(d_0,D_{c_0},D_{d_0})\ell+\delta^2\bar{p}_1(D_{c_0})\frac{c}{1-\delta^2}+\delta^2\bar{p}_1(D_{d_0})\ell\\
&=\overline{\alpha}(\delta)\frac{c}{1-\delta^2}+\{1-\overline{\alpha}(\delta)\}\ell\\
&\qquad+\delta^2\bar{p}_1(D_{c_0},D_{d_0})[\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}\cdot\frac{c}{1-\delta^2}+\Big(1-\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}\Big)\ell]\\
&=U_0(c_0;\overline{\alpha}(\delta))+\delta^2\bar{p}_1(D_{c_0},D_{d_0})\cdot U_0\Big(c_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}\Big).
\end{aligned}
$$

Hence

$$
\Delta U(C_{c_0},c_0;\bar{p}_1)=U(C_{c_0};\bar{p}_1)-U(c_0;\bar{p}_1)=\delta^2\bar{p}_1(D_{c_0},D_{d_0})\cdot U_0\Big(c_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}\Big).
$$

That is, the payoff difference between C_{c_0} and c_0 is the continuation payoff in the second period of a partnership in which (C,C) is not established in the first period, or with the partners using a D_z-strategy for some $z\in\{c_0,d_0\}$. Similarly,

$$
\begin{aligned}
L(C_{c_0};\bar{p}_1)&=\overline{\alpha}(\delta)\frac{1}{1-\delta^2}+\{1-\overline{\alpha}(\delta)\}\cdot 1+\delta^2\bar{p}_1(D_{c_0})\cdot\frac{1}{1-\delta^2}+\delta^2\bar{p}_1(D_{d_0})\cdot 1\\
&=L_0(c_0;\overline{\alpha}(\delta))+\delta^2\bar{p}_1(D_{c_0},D_{d_0})\cdot L_0\Big(c_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}\Big),\\
\Rightarrow \Delta L(C_{c_0},c_0;\bar{p}_1)&=\delta^2\bar{p}_1(D_{c_0},D_{d_0})\cdot L_0\Big(c_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}\Big).
\end{aligned}
$$

If $\bar{p}_1(D_{c_0},D_{d_0})=0$, then $\Delta U(C_{c_0},c_0;\bar{p}_1)=0$ and $\Delta L(C_{c_0},c_0;\bar{p}_1)=0$, so that $v(C_{c_0};\bar{p}_1)=v(c_0;\bar{p}_1)$ holds. If $\bar{p}_1(D_{c_0},D_{d_0})>0$, from (7) or (8),

$$
\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}=\overline{\alpha}(\delta)
$$

so that the **average payoff difference** between the C_{c_0}-strategy and the c_0-strategy is

$$
\frac{\Delta U(C_{c_0},c_0;\bar{p}_1)}{\Delta L(C_{c_0},c_0;\bar{p}_1)}=\frac{U_0\Big(c_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}\Big)}{L_0\Big(c_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}\Big)}=v_0\Big(c_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0},D_{d_0})}\Big)=v_0(c_0;\overline{\alpha}(\delta)).
$$

This means that the extra continuation average payoff that the C_{c_0}-strategy gets by its tolerance is the same as the average payoff of going back to the matching pool. By Lemma 1 (i), we have

$$
v(C_{c_0};\bar{p}_1)=v(c_0;\bar{p}_1).
$$

Next, consider C_{d_0}-strategy. Again, from Table 3,

$$U(C_{d_0};\bar{p}_1) = U_0(c_0;\overline{\alpha}(\delta)) + \delta^2\bar{p}_1(D_{c_0}, D_{d_0}) \cdot U_0\Big(d_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0}, D_{d_0})}\Big);$$
$$L(C_{d_0};\bar{p}_1) = L_0(c_0;\overline{\alpha}(\delta)) + \delta^2\bar{p}_1(D_{c_0}, D_{d_0}) L_0\Big(d_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0}, D_{d_0})}\Big).$$

The payoff difference between C_{d_0} and c_0 is the continuation payoff difference after histories in which (C, C) has not been established, or with the partners using a D_z-strategy for some $z \in \{c_0, d_0\}$. If $\bar{p}_1(D_{c_0}, D_{d_0}) = 0$, then $v(C_{d_0};\bar{p}_1) = v(c_0;\bar{p}_1)$ holds. Otherwise,

$$\frac{\Delta U(C_{d_0}, c_0;\bar{p}_1)}{\Delta L(C_{d_0}, c_0;\bar{p}_1)} = v_0\Big(d_0;\frac{\bar{p}_1(D_{c_0})}{\bar{p}_1(D_{c_0}, D_{d_0})}\Big) = v_0(c_0;\overline{\alpha}(\delta)).$$

By Lemma 1 (i), we have $v(C_{d_0};\bar{p}_1) = v(c_0;\bar{p}_1)$. □

Step 3: For each $s \in \{D_{c_0}, D_{d_0}\}$, $v(s;\bar{p}_1) = v(d_0;\bar{p}_1)$.

Proof of Step 3: Analogous to the proof of Step 2.

Therefore, all strategies in $supp(\bar{p}_1)$ are payoff equivalent to one another and to the fundamentally asymmetric equilibrium $p_{c_0 d_0}(\delta)$.

Step 4: Consider any (on-path) partnership history such that the previous action combination was (C, C). A one-step deviation from any strategy in $supp(\bar{p}_1)$ involves either (Stay, D) or Leave. None of these gets a higher continuation payoff than that of the equilibrium strategies.

Proof of Step 4: Let v^* be the common average payoff of the equilibrium strategies starting in the matching pool;

$$v^* = v(s;\bar{p}_1),\ \forall s \in supp(\bar{p}_1).$$

Hence $V(s;\bar{p}_1) = v^*/(1-\delta)$. The assumption that (C, C) is observed implies that the future play path of this partnership must be (Stay, Stay), (C, C), (Stay, Stay), (C, C), $\ldots$ as long as both partners live (with probability δ^2 each period). However, with probability $\delta(1-\delta)$, a player survives the partner and goes back to the matching pool in the next period. Hence the (non-averaged) continuation payoff of any of the equilibrium strategies (measured at the end of a period right after the observation of

(C,C) in the Prisoner's Dilemma) is

$$
\begin{aligned}
&\delta^2[\frac{c}{1-\delta^2} + \{\delta(1-\delta) + \delta^3(1-\delta) + \cdots\} \cdot \frac{v^*}{1-\delta}] + \delta(1-\delta)\frac{v^*}{1-\delta} \\
=&\delta^2[\frac{c}{1-\delta^2} + \frac{\delta(1-\delta)}{1-\delta^2} \cdot \frac{v^*}{1-\delta}] + \delta(1-\delta)\frac{v^*}{1-\delta}. \qquad (9)
\end{aligned}
$$

By contrast, any one-step deviation to (Stay, D) gives g in the next period (if the partnership continues), but the match is terminated by the partner (who follows some strategy in $supp(\bar{p}_1)$). Hence the continuation payoff of such a one-step deviation strategy is

$$\delta^2[g + \delta \cdot \frac{v^*}{1-\delta}] + \delta(1-\delta)\frac{v^*}{1-\delta},$$

and this is strictly less than (9) by Lemma 4.

Any one-step deviation to leave after (C,C) but conforms to an equilibrium strategy s gives the (expected) continuation payoff of $\delta V(s;\bar{p}_T) = \delta \cdot \frac{v^*}{1-\delta}$. Since $v^* = v_0(c_0;\overline{\alpha}(\delta)) < g$,

$$\delta \cdot \frac{v^*}{1-\delta} < \delta^2[g + \delta \cdot \frac{v^*}{1-\delta}] + \delta(1-\delta)\frac{v^*}{1-\delta}.$$

Hence this kind of one-step deviation is worse than (9) as well. □

Notice that after an action profile $(a,a') \neq (C,C)$ in the first period of a match, all of Leave, (Stay, C) and (Stay, D) are on-path actions. Therefore, it remains to consider $(a,a') \neq (C,C)$ in the second period of a match.

Step 5: Consider any on-path partnership history such that the second period action combination was (C,D), (D,C), or (D,D). A one-step deviation from some strategy in $supp(\bar{p}_1)$ is either (Stay, C) or (Stay, D). However, the partner (who follows some strategy in $supp(\bar{p}_1)$) terminates the match after $(a,a') \neq (C,C)$, and hence these deviations all result in the same continuation payoff $\delta \cdot \frac{v^*}{1-\delta}$ which is also the continuation payoff of the equilibrium strategies.

This completes the proof of Proposition 1. *Q.E.D.*

Intuitively, the construction of one-period tolerant equilibria is by the **internalization** of the fundamentally asymmetric equilibrium, $p_{c_0d_0}(\delta)$, when its play path leads them to the brink of dissolving

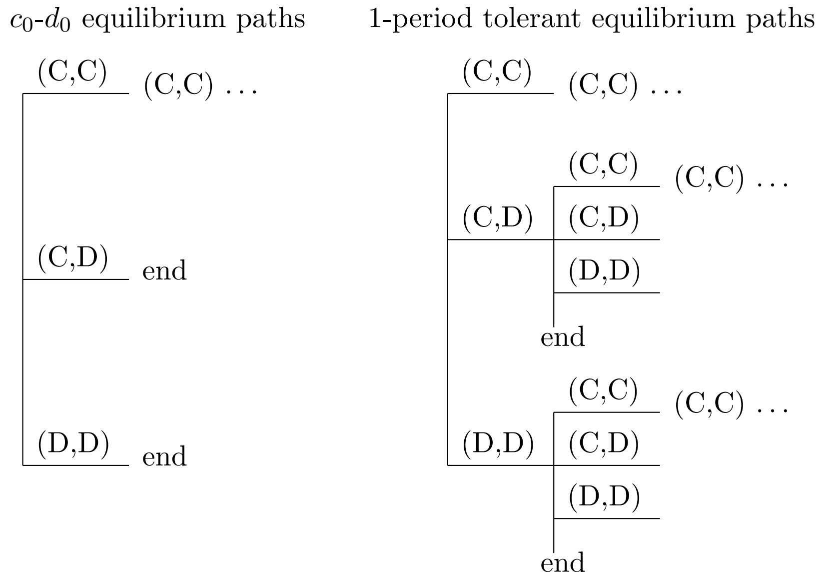

Figure 3: Internalization of the matching pool distribution

(see Figure 3). Consider the moment when a player is about to leave a partnership in the $p_{c_0 d_0}(\delta)$-equilibrium. If the probability distribution over the possible continuation strategies of the current opponent is the same as the $p_{c_0 d_0}(\delta)$-distribution, then a player is indifferent between staying and going back to the matching pool.[18] Moreover, in this case, a player can play either the c_0-strategy or the d_0-strategy as the continuation strategy as well. Figure 3 illustrates that these strategies generate all two-period action combination sequences consistent with the experimental observation (ii) that (C,C) implies (Stay, Stay) and (C,C). The next corollary is straightforward.

Corollary 1 *(Behavioral diversity for two periods) Let*

$$A^2_{co} := \{[(a_1, a'_1), (a_2, a'_2)] \in \{C,D\}^2 \times \{C,D\}^2 \mid (a_1, a'_1) = (C,C) \Rightarrow (a_2, a'_2) = (C,C)\}$$

be the set of two-period "conditionally cooperative" PD action combination sequences. For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$*, any* $\bar{p}_1(\delta)$ *satisfying (7) with some* $\beta, \gamma \in (0,1)$ *generates all action combination sequences in* A^2_{co} *with a positive probability.*

[18] This part of the logic is similar to the construction of the Co-De-indifferent equilibrium by Lei et al. (2018). However our working paper (Fujiwara-Greve and Okuno-Fujiwara, 2016) precedes their work. Moreover, we allow any length of memory, and therefore we are able to capture behavioral diversity.

3.2 General tolerant equilibria

The construction of the class of 1-period tolerant strategies can be generalized by prolonging the initial tolerant phase with a planned action sequence (see also Figure 5 in Section 3.5).

Definition 7 For any $k = 0, 1, 2, \ldots$, any k-period sequence[19] $\mathbf{X} = (X_1, X_2, \ldots, X_k) \in \{C, D\}^k$, and any $z \in \{c_0, d_0\}$, the *k-period tolerant strategy*, denoted $\mathbf{X}_z$, is a match-independent strategy such that, for any period $t = 1, 2, \ldots$ in a partnership,
if $t \leqq k$ (Tolerant phase): when (C, C) is observed in the previous period[20], play the c_0-strategy as the continuation strategy, and otherwise play $X_t \in \{C, D\}$ and stay for any observation;
if $t = k+1$ (Commitment phase): when (C, C) is observed in the previous period, play the c_0-strategy as the continuation strategy, and otherwise play the z-strategy as the continuation strategy.

This class includes the $\mathbf{X}_z$-strategy with $\mathbf{X} = \overbrace{D \cdots D}^{k \text{ times}}$ and $z = c_0$ (we write this strategy as $D^k_{c_0}$), which is a tolerant version of the k-period *trust-building strategy* in Greve-Okuno.

The set of k-period tolerant strategies is denoted by

$$\tilde{S}_k = \{\mathbf{X}_z \in \mathbf{S} \mid \mathbf{X} \in \{C, D\}^k,\ z \in \{c_0, d_0\}\}.$$

For each $k = 0, 1, 2, \ldots$, the set of all tolerant strategies with k-period or longer tolerant phase is denoted by $\tilde{S}^\infty_k := \cup^\infty_{j=k} \tilde{S}_j$.

For notational convenience, define the set of all "C-start" and all "D-start" tolerant strategies:

$$C_+ := \{\mathbf{X}_z \in \tilde{S}^\infty_0 \mid X_1 = C\}$$

$$D_+ := \{\mathbf{X}_z \in \tilde{S}^\infty_0 \mid X_1 = D\}.$$

For example, $C_+ \cap \tilde{S}^1_0 = \{c_0, C_{c_0}, C_{d_0}\}$. Similarly, for each for each $k = 0, 1, 2, \ldots$ and any k-period action sequence $\mathbf{X} \in \{C, D\}^k$, define the set of "$\mathbf{X}C$-start" tolerant strategies and "$\mathbf{X}D$-start" tolerant

[19] As a convention, let $\{C, D\}^0 = \emptyset$. Hence the c_0- and the d_0-strategies are included as degenerate tolerant strategies.
[20] If $t = 1$, assume that this is not the case.

strategies[21]:

$$\mathbf{X}C_+ := \{\mathbf{X}'_s \in \tilde{S}_k^\infty \mid \mathbf{X}' = \mathbf{X},\ s \in C_+\}$$

$$\mathbf{X}D_+ := \{\mathbf{X}'_s \in \tilde{S}_k^\infty \mid \mathbf{X}' = \mathbf{X},\ s \in D_+\}.$$

(For $k = 0$, we use the convention $\emptyset C_+ = C_+$ and $\emptyset D_+ = D_+$.) For example, $CC_+ \subset \tilde{S}_1^\infty$ and $CC_+ \cap \tilde{S}_1^2 = \{C_{c_0}, CC_{c_0}, CC_{d_0}\}$.

Note that we can decompose the set $\mathbf{X}C_+$ and $\mathbf{X}D_+$ indefinitely. For example,

$$\begin{aligned}\mathbf{X}C_+ &= \{\mathbf{X}_{c_0}\} \cup \mathbf{X}CC_+ \cup \mathbf{X}CD_+ \\ &= \{\mathbf{X}_{c_0}\} \cup [\{\mathbf{X}C_{c_0}\} \cup \mathbf{X}CCC_+ \cup \mathbf{X}CCD_+] \cup [\{\mathbf{X}C_{d_0}\} \cup \mathbf{X}CDC_+ \cup \mathbf{X}CDD_+]\end{aligned}$$

and so on. This means that if the support of some p does not contain $\mathbf{X}C_+$, then it does not contain any of $\mathbf{X}CC_+, \mathbf{X}CD_+, \ldots$ which are the set of longer tolerant strategies with the same initial action sequence $(X_1, \ldots, X_k, C)$ (when $\mathbf{X} = (X_1, \ldots, X_k)$).

Proposition 2 *(Tolerant Equilibria) Fix any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$. For each $T = 0, 1, 2, \ldots,$ define*

$$P_T(\delta) := \Big\{\bar{p}_T \in \mathcal{P}(\tilde{S}_0^T) \mid \ \forall k = 0, 1, 2, \ldots, T,\ \forall \mathbf{X} \in \{C, D\}^k, \\ \bar{p}_T\big(\mathbf{X}C_+ \cup \mathbf{X}D_+\big) > 0 \Rightarrow \frac{\bar{p}_T\big(\mathbf{X}C_+\big)}{\bar{p}_T\big(\mathbf{X}C_+ \cup \mathbf{X}D_+\big)} = \overline{\alpha}(\delta)\Big\}. \tag{10}$$

Then, any distribution in $\bar{p}_T \in P_T(\delta)$ is a Nash equilibrium and is payoff-equivalent to the bimorphic $p_{c_0 d_0}(\delta)$-equilibrium.

Proof. See Appendix.

The strategy distributions in $P_T(\delta)$ have a "branching" structure such that if a strategy in $\mathbf{X}C_+$ or $\mathbf{X}D_+$ exists in the support, then the relative ratio of these groups must be the C-D ratio (so that the other group also exists in the support), but it is possible that, for some $\mathbf{X}$, none of $\mathbf{X}C_+$-start and $\mathbf{X}D_+$-start tolerant strategies exist in the support.

[21] Given an action plan $\mathbf{X} \in \{C, D\}^k$ and a tolerant strategy $s = \mathbf{Y}_z$, the "concatenated" strategy $\mathbf{X}_s$ is also a tolerant strategy of the form $\mathbf{X}\mathbf{Y}_z$.

3.3 Behavioral diversity

If a strategy distribution in $P_T(\delta)$ has the full-support in the sense that all of 0 to T-period tolerant strategies exist in its support, it generates all Prisoner's Dilemma action profile sequences for $T+1$ periods which are consistent with the experimental observation (ii) that (C, C) implies (Stay, Stay) and (C, C). To state this formally, for each $T = 0, 1, 2, \ldots$, define the set of "full-support" tolerant equilibria:

$$P_T^\circ(\delta) := \Big\{ \bar{p}_T \in P_T(\delta) \mid \forall k = 0, 1, 2, \ldots, T,\ \forall \mathbf{X} \in \{C, D\}^k,\ \bar{p}_T(\mathbf{X}_z) > 0, \ \ \forall z \in \{c_0, d_0\} \Big\}.$$

For any $n = 1, 2, \ldots$, let a class of n-period action profile sequences with the property that the establishment of (C, C) implies long-term cooperation:

$$\begin{aligned} A_{co}^n := \ & \Big\{ ((a_1, a_1'), \ldots, (a_n, a_n')) \in [\{C, D\}^2]^n \mid \forall m = 1, 2, \ldots, n-1, \\ & \qquad (a_m, a_m') = (C, C) \Rightarrow (a_{m+1}, a_{m+1}') = (C, C) \Big\}. \end{aligned}$$

Corollary 2 *(Behavioral Diversity) For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$ and any $T = 1, 2, \ldots$, each distribution in $\bar{p}_T \in P_T^\circ(\delta)$ generates all $T+1$-period action combination sequences in A_{co}^{T+1} with a positive probability.*

The proof is straightforward from the definition of the tolerant strategies and the requirement that any k-period tolerant strategy for $k = 0, 1, \ldots, T$ must exist in $\bar{p}_T$ (recall also Figure 3).

3.4 Efficiency

By construction, all equilibria in $P_T(\delta)$ for all $T = 0, 1, 2, \ldots$ are payoff-equivalent to one another, for a given $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$. Fujiwara-Greve et al. (2015) showed that when the "stake" $(g - c)$ of the Prisoner's Dilemma is not too large, the fundamentally asymmetric equilibrium is more efficient than any trust-building/gradual cooperation equilibrium.

Definition 8 For any $k = 0, 1, 2, \ldots$, the *k-period trust-building strategy*, denoted c_k, is a match-independent strategy such that, for any period $t = 1, 2, \ldots$ in a partnership, if $t \leqq k$ (Trust-building phase): play D and stay if and only if (D, D) is observed;

if $t \geqq k+1$ (Cooperation phase): play C and stay if and only if (C, C) is observed.

Remark 1 *(Proposition 4 of Fujiwara-Greve et al., 2015) For any* (g, c, d, ℓ) *such that* $g - c < (c - d)^2/(c - \ell)$*, there exists* $\hat{\delta} \in [\underline{\delta}_{c_0 d_0}, 1)$ *such that for any* $\delta \in (\hat{\delta}, 1)$*, any Nash equilibrium distribution* $q \in \mathcal{P}(\{c_0, c_1, c_2, \dots\})$*, and any* $s \in \{c_0, d_0\}$*,*

$$v(s; p_{c_0 d_0}(\delta)) \geqq v(s'; q), \ \forall s' \in supp(q),$$

and the strict inequality holds for any Nash equilibrium $q \in \mathcal{P}(\{c_1, c_2, \dots\})$ *and any* $s' \in supp(q)$*.*

The idea is that, under the small stake condition, occasional exploitation of c_0-players by d_0-players is not so costly as compared to **all players suffering** from mutual defection at the beginning of **every match**. Therefore, under the same payoff condition and δ's, **all tolerant equilibria are more efficient** than any trust-building equilibrium.

3.5 Evolutionary Stability

Since the model is a population game (Sandholm, 2010), it is desirable to investigate evolutionary stability of the tolerant equilibria. However, the neutral stability concept in Greve-Okuno cannot be satisfied by any tolerant equilibrium, just like the fundamentally asymmetric c_0-d_0 equilibrium does not satisfy it. Greve-Okuno required that **all** equilibrium strategies must perform weakly better than any pure-strategy mutant. To distinguish from a more standard way to compare the mean payoff of strategy distributions (see Definition 12), we rename their stability concept as S-Neutral Stability[22], where S stands for symmetric/single-strategy mutants, and also make the set of possible mutants explicit, which is $\mathbf{S}$ in this concept.

Definition 9 *(Greve-Okuno)* A stationary strategy distribution in the matching pool $p^* \in \mathcal{P}(\mathbf{S})$ satisfies *S-Neutral Stability with respect to* $\mathbf{S}$ (denoted S-NS($\mathbf{S}$)) if, for any $s' \in \mathbf{S}$, there exists $\bar{\epsilon} \in (0, 1)$

[22] Izquierdo et al. (2018) call this concept Fujiwara NSD.

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

$$v(s; (1-\epsilon)p^* + \epsilon \cdot s') \geqq v(s'; (1-\epsilon)p^* + \epsilon \cdot s'). \tag{11}$$

When there is on-path separation among the equilibrium strategies, there is a **secret-handshake** strategy (Robson, 1990, Matsui, 1991) which imitates an equilibrium strategy until when the latter chooses to leave, stay instead and cooperate afterwards. If such a secret-handshake strategy enters the population with a positive measure, the imitated strategy performs strictly worse than the secret-handshake mutants. Hence no tolerant equilibrium satisfies S-NS($\mathbf{S}$), or even if we restrict the set of symmetric-strategy mutants to $\tilde{S}_0^{\infty}$.

Remark 2 *For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, any $T = 0, 1, 2, \ldots$, and any $\bar{p}_T \in P_T(\delta)$, $\bar{p}_T$ does not satisfy S-NS($\mathbf{S}$).*

Proof of Remark 2. (Straightforward and can be omitted.) Fix an arbitrary $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, any $T = 0, 1, 2, \ldots$ and any $\bar{p}_T \in P_T(\delta)$. Among the equilibrium pure strategies, consider the most-tolerant, always-D strategy $s = D_{d_0}^k$. That is, take $s \in \tilde{S}_k \cap supp(\bar{p}_T)$ such that there is no $\hat{s} = D_{d_0}^m \in \tilde{S}_m \cap \mathrm{supp}(\bar{p}_T)$ with $m > k$. (For $T = 0$, s must be the d_0-strategy. For $T = 1$, s is either d_0 or D_{d_0} depending on the support of $\bar{p}_1$.) Then there exists a "secret-handshake" mutant strategy $s' = D_{c_0}^{k+1}$ which imitates s but stays in the partnership in $k+1$-th period and starts cooperating in $k+2$-th period. This strategy has the same play path as that of the s-strategy against any equilibrium strategy, and can establish long-term cooperation with another s'-strategy. Therefore, for any $\epsilon \in (0, 1)$,

$$\frac{\Delta U(s', s; (1-\epsilon)\bar{p}_T + \epsilon \cdot s')}{\Delta L(s', s; (1-\epsilon)\bar{p}_T + \epsilon \cdot s')} = c > v(s; (1-\epsilon)\bar{p}_T + \epsilon \cdot s'),$$

and $\Delta L(s', s; (1-\epsilon)\bar{p}_T + \epsilon \cdot s') > 0$. By Lemma 1 (i), for any $\epsilon \in (0, 1)$,

$$v(s; (1-\epsilon)\bar{p}_T + \epsilon \cdot s') < v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot s'). \qquad \textit{Q.E.D.}$$

In other words, the tolerant equilibria are not robust against coordinated mutants using a secret-handshake strategy, which exists among the tolerant strategies.

The S-NS($\mathbf{S}$) concept by Greve-Okuno was defined mostly for the monomorphic distribution analysis where both the equilibrium-strategy distribution p^* and the potential mutant-strategy distributions consist of a single pure-strategy. However, if p^* has many pure strategies in the support, after mutants perturb the matching-pool strategy distribution, incumbent strategies differ in the post-entry payoffs. Hence, requiring **each** equilibrium strategy to beat all mutants is very strong. A more standard measurement of the fitness/performance of a strategy distribution is to look at the mean payoff over the relevant strategy distribution.

Also, we should include polymorphic-strategy distributions as potential mutant distributions. However, it is not fruitful to consider any distribution in $\mathcal{P}(\mathbf{S})$ from Remark 2. Thus we define a stability concept **with respect to** a set of potential (polymorphic) mutant-strategy distributions, and investigate a "sufficient" set of mutant distributions which cannot de-stabilize the tolerant equilibria. To avoid measure theoretic complications which do not give us new economic insights, we restrict our attention to countable support distributions. Let $\mathcal{Q} \subset \mathcal{P}(\mathbf{S})$ be the set of (match-independent) strategy distributions with a countable support.

Definition 10 Given $M \subset \mathcal{Q}$, a stationary strategy distribution in the matching pool $p^* \in \mathcal{Q}$ is *Mean Stable with respect to the mutants from M* (denoted MS(M)) if, for any $q \in M$, there exists $\bar{\epsilon} \in (0,1)$ such that for any $\epsilon \in (0, \bar{\epsilon})$,

$$\sum_{s \in supp(p*)} p^*(s) v(s; (1-\epsilon)p^* + \epsilon \cdot q) > \sum_{s' \in supp(q)} q(s') v(s'; (1-\epsilon)p^* + \epsilon \cdot q). \tag{12}$$

Since we need to restrict the potential mutant distribution set M, we might as well require the strict inequality. Note that the larger the M is, the stronger the stability is. If (12) holds, some kind of monotone/mean dynamic (see e.g., Samuelson, 1997 and Sandholm, 2010) should support that any mutant strategy distribution will be expelled, when mutant distributions are generated within M.

If $M = Q$ and (12) is the weak inequality, Definition 10 is the "matching-pool distribution" version of the *Izquierdo & van Veelen Neutrally Stable Distribution* in Izquierdo et al. (2018). Note also that

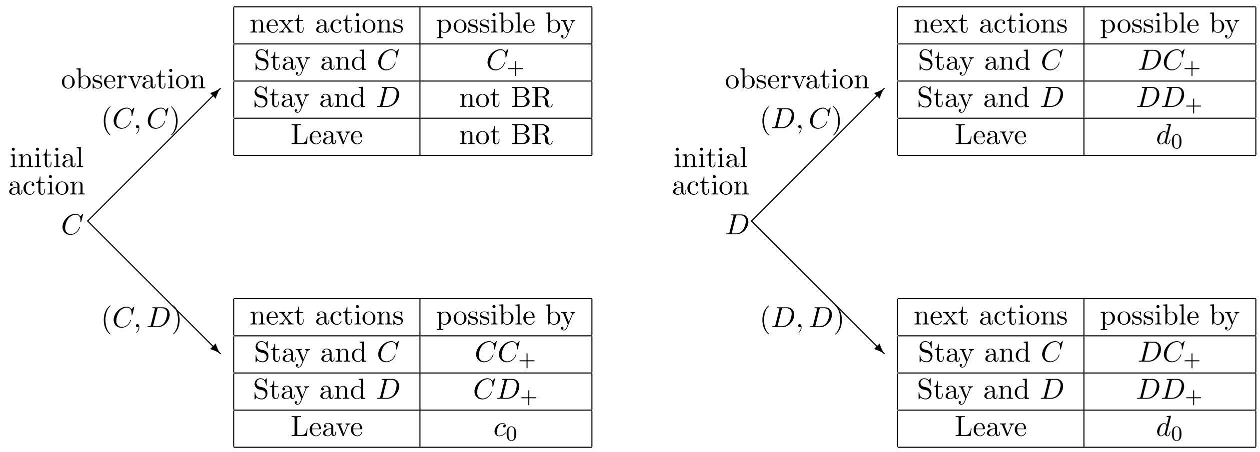

Figure 4: Feasible action sequences over two periods by tolerant strategies

for any $p^*, q \in \mathcal{Q}$, the "strategy-wise" comparison of the average payoffs as in (11),

$$\forall s' \in supp(q),\ \forall s \in supp(p^*),\ v(s; (1-\epsilon)p^* + \epsilon \cdot s') > v(s'; (1-\epsilon)p^* + \epsilon \cdot s'),$$

may not imply (12), because, when q is polymorphic,

$$\sum_{s' \in supp(q)} q(s')v(s'; (1-\epsilon)p^* + \epsilon \cdot q) \neq \sum_{s' \in supp(q)} q(s')v(s'; (1-\epsilon)p^* + \epsilon \cdot s').$$

This means that the Mean Stability is fundamentally different from the stability considered in Greve-Okuno.

In order to establish a Mean Stability result of **all** tolerant equilibria, we focus on mutants using only tolerant strategies as well. A justification for this focus is that the tolerant strategies generate **all relevant** play paths in the society. By behavioral diversity (Corollary 2), the infeasible play paths have the property [Stay and D] or Leave after observing (C, C) (see Figure 4). Any strategy which generates such a play path is a one-step deviation strategy and thus is not a best response. In other words, tolerant strategies are the candidates for **equilibrium entrants**.[23]

Another justification of our focus on the tolerant strategy mutants is that, among the tolerant strategies, the **explicit payoff comparison is possible**. In general, it is very difficult to explicitly

[23]For a related concept, see Swinkels (1992).

compare the average payoffs of two (extensive-form) strategies such as

$$v(s;p) = \frac{U(s;p)}{L(s;p)} = \frac{\sum_{s' \in supp(p)} p(s')U(s,s')}{\sum_{s' \in supp(p)} p(s')L(s,s')}, \quad v(\hat{s};p) = \frac{U(\hat{s};p)}{L(\hat{s};p)} = \frac{\sum_{s' \in supp(p)} p(s')U(\hat{s},s')}{\sum_{s' \in supp(p)} p(s')L(\hat{s},s')},$$

because both the denominator and the numerator can be different, due to the endogenous lengths of partnerships. By contrast, the structure of the tolerant strategies is so organized that for any pair of tolerant strategies s and s^* with the same initial action plan for some periods and s^* has a shorter tolerant phase than that of s (then we call that s^* is an *induced* strategy of s),

$$v_0(z;q) \leqq \frac{U(s;p) - U(s^*;p)}{L(s;p) - L(s^*;p)} \leqq v_0(z';q')$$

for some $z, z' \in \{c_0, d_0\}$ and some $q, q' \in [0,1]$. Then Lemma 1 can be invoked to compare $v(s;p)$ and $v(s^*;p)$ explicitly. (See the proof of Lemma 5 below.)

Definition 11 For any $k = 1, 2, \ldots$ any k-period action sequence $\mathbf{X} = (X_1, \ldots, X_k) \in \{C, D\}^k$, any $z \in \{c_0, d_0\}$, and any $m = 0, 1, 2, \ldots, k-1$, the *induced m-period tolerant strategy* of the (k-period tolerant) $\mathbf{X}_z$-strategy is an m-period tolerant strategy $\mathbf{X}'_{x_{m+1}}$ such that the planned action sequence for initial m periods is the same, $\mathbf{X}' = (X_1, \ldots, X_m)$, and the commitment continuation strategy in $m+1$-th period of a match is

$$x_{m+1} = \begin{cases} c_0 & \text{if } X_{m+1} = C \\ d_0 & \text{if } X_{m+1} = D. \end{cases}$$

Figure 5 illustrates a k-period tolerant strategy and its induced $(k-1)$-period tolerant strategy.[24] They behave the same way **for k-periods** but the induced $(k-1)$-period tolerant strategy commits to one of the c_0- or the d_0-strategy in the k-th period, depending on $X_k = C$ or D. In general, a k-period tolerant strategy and its induced m-period tolerant strategy (where $k > m$) behave the same way for $(m+1)$-period. This simplifies the average payoff comparison.

Next, we give an important lemma which shows **External Stability**[25] (if mutant distributions are within some set of tolerant strategies): for any tolerant strategy mutant s', there is an equilibrium

[24] As in the case of D_z-strategies, those which do not choose C at some point can be vacuously illustrated as in Figure 5 as well.

[25] This property is similar to von Neumann-Morgenstern's External Stability which is a part of their notion of Stable Set (von Neumann and Morgenstern, 1944).

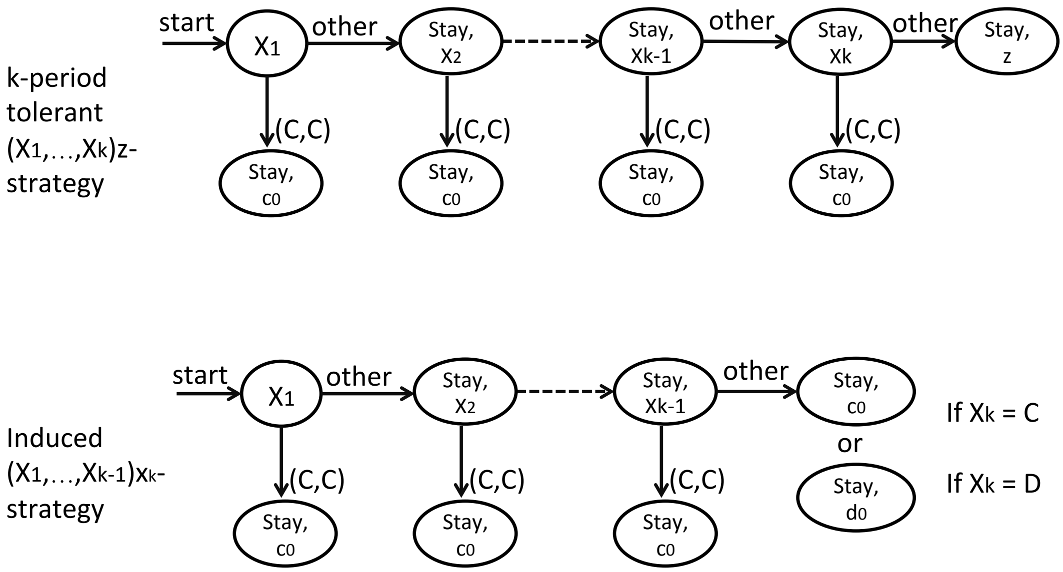

Figure 5: k-period tolerant strategy and its induced $(k-1)$-period tolerant strategy

strategy s^* which performs strictly better than s' does in the post-entry distribution. To show this, we partition the mutant (tolerant) strategies using the induced-strategy structure. For each $T = 0, 1, 2, \ldots$, and each (T-period tolerant) $s_T \in \tilde{S}_T$, define

$$S(s_T) := \{s' \in \tilde{S}^{\infty}_{T+1} \mid s_T \text{ is the induced T period tolerant strategy of } s'\}.$$

Then

$$S(s_T) \cap S(\hat{s}_T) = \emptyset,\ \forall s_T \neq \hat{s}_T \in \tilde{S}_T,\ \cup_{s_T \in \tilde{S}_T} S(s_T) = \tilde{S}^{\infty}_{T+1}.$$

Lemma 5 *(External Stability) For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, and each $T = 0, 1, 2, \ldots$, define*

$$M_T(\delta) := \Big\{q \in \mathcal{P}(\tilde{S}^{\infty}_0) \mid supp(q) \cap \tilde{S}^{\infty}_{T+1} \neq \emptyset,\ \sup_{\substack{\mathbf{X} \in \{C,D\}^{T+k},\ k=1,2,\ldots \\ q(\mathbf{X}C_+ \cup \mathbf{X}D_+) > 0}} \frac{q\big(\mathbf{X}C_+\big)}{q\big(\mathbf{X}C_+ \cup \mathbf{X}D_+\big)} < \overline{\alpha}(\delta)\Big\}.$$

Then for any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$ and any $T = 1, 2, \ldots$, each full-support $\bar{p}_T \in P^{\circ}_T$ satisfies the following: for any $q \in M_T(\delta)$, there exists $\bar{\epsilon}_T \in (0, 1)$ such that, for any $\epsilon \in (0, \bar{\epsilon}_T)$ and any $s_T \in \tilde{S}_T$,

$$\forall s' \in S(s_T) \cap supp(q),\ v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) > v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q). \tag{13}$$

Proof. See Appendix.

Lemma 5 is in fact stronger than External Stability because we know that for any (mutant strategy) $s' \in \tilde{S}^{\infty}_{T+1}$, it is outperformed by its induced T-period tolerant strategy, which is in the support of $\bar{p}_T$.

The sufficient set of mutant strategies, $M_T(\delta)$, can be interpreted as "'dispersed distributions" such that mutants do not concentrate heavily on some $\mathbf{X}_{c_0}$-strategy. $M_T(\delta)$ requires only that tolerant strategies whose tolerant phase last $T+1$ **periods or longer** emerge in a dispersed way. Since there are infinitely many strategies even within the tolerant strategy class, unlike Kandori et al. (1993), there is no guarantee that in the long-run, mutants can coordinate on a particular strategy. Instead, it is natural that mutant strategy distributions have a wide support. ($M_T(\delta)$ includes q's with full-support.)

$M_T(\delta)$ is nearly the necessary and sufficient set for external stability. From Proposition 2, if the relative ratio of mutant strategies is exactly the C-D ratio, the post-entry distribution becomes a new Nash equilibrium. From Remark 2, if a mutant distribution concentrates on some secret-handshake $\mathbf{X}_{c_0}$-strategy, they can invade a tolerant equilibrium.

The evolutionary stability result of **all** tolerant equilibria is as follows. To construct a non-empty interior, we consider upper bounds to the tolerant phase in the society, $m = T+1, T+2, \ldots$. For each $m < \infty$, we can use the ordinary, finite-dimensional Euclidean space topology on $\mathcal{P}(\tilde{S}^m_0)$.

Proposition 3 *For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, any $T = 0, 1, 2, \ldots$, any $m = T+1, T+2, \ldots$, and any full support $\bar{p}_T \in P^{\circ}_T$, there exists $M \subset M_T(\delta) \cap \mathcal{P}(\tilde{S}^m_0)$ with a non-empty interior such that $\bar{p}_T$ is MS(M).*

Proof. See Appendix.

The idea of the proof is as follows. There exists (in fact many) $q \in M_T(\delta) \cap int(\mathcal{P}(\tilde{S}^m_0))$ such that

$$\forall k = 0, 1, \ldots, T-1,\ \forall\ s_k \in \tilde{S}_k,\ q(s_k) = \bar{p}_T(s_k),$$

$$\forall s_T \in \tilde{S}_T,\ q(s_T) + \sum_{s' \in S(s_T)} q(s') = \bar{p}_T(s_T).$$

By External Stability of Lemma 5, for sufficiently small ϵ's, and each $s_T \in \tilde{S}_T$,

$$
\begin{aligned}
& \sum_{s' \in S(s_T) \cap supp(q)} q(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) + q(s_T)v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
< & \sum_{s' \in S(s_T) \cap supp(q)} q(s')v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) + q(s_T)v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
= & [q(s_T) + \sum_{s' \in S(s_T) \cap supp(q)} q(s')] \cdot v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
= & \bar{p}_T(s_T) \cdot v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \qquad \text{from the definition of q.}
\end{aligned}
$$

Also by the definition of q, for any $k = 0, 1, \ldots, T-1$ and $s_k \in \tilde{S}_k$,

$$
q(s_k)v(s_k; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) = \bar{p}_T(s_k)v(s_k; (1-\epsilon)\bar{p}_T + \epsilon \cdot q).
$$

Hence

$$
\begin{aligned}
& \sum_{s' \in supp(q)} q(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
= & \sum_{k=0}^{T-1} \sum_{s_k \in \tilde{S}_k} q(s_k)v(s_k; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
& + \sum_{s_T \in \tilde{S}_T} q(s_T)v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) + \sum_{s_T \in \tilde{S}_T} \sum_{s' \in S(s_T) \cap supp(q)} q(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
< & \sum_{k=0}^{T-1} \sum_{s_k \in \tilde{S}_k} \bar{p}_T(s_k)v(s_k; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) + \sum_{s_T \in \tilde{S}_T} \bar{p}_T(s_T) \cdot v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
= & \sum_{s \in \tilde{S}_0^T} \bar{p}_T(s_T)v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q).
\end{aligned}
$$

Since we constructed q in the interior of $\mathcal{P}(\tilde{S}_0^m)$, we can find an open ball around it.

However, this is only an example of the existence of M for MS(M), and unfortunately we do not have a result closer to the characterization of the sufficient set M for Mean Stability. The difficulty lies in the non-linear average payoff structure and the vast variety of strategies in the VSRPD.

For the simplest case of $T = 0$, our companion paper Fujiwara-Greve and Okuno-Fujiwara (2019) shows that in fact the set very similar to $M_0(\delta)$ makes the fundamentally asymmetric equilibrium Neutrally Stable, which is slightly weaker than Mean Stability.

Definition 12 Given $M \subset \mathcal{Q}$, a stationary strategy distribution in the matching pool $p^* \in \mathcal{Q}$ satisfies *Neutral Stability with respect to the mutants from* M (denoted NS(M)) if, for any $q \in M$, there exists $\overline{\epsilon} \in (0,1)$ such that for any $\epsilon \in (0, \overline{\epsilon})$,

$$\sum_{s \in supp(p*)} p^*(s)v(s;(1-\epsilon)p^* + \epsilon \cdot q) \geqq \sum_{s' \in supp(q)} q(s')v(s';(1-\epsilon)p^* + \epsilon \cdot q). \tag{14}$$

Remark 3 *(Fujiwara-Greve and Okuno-Fujiwara (2019)) For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, the bimorphic Nash equilibrium $p_{c_0 d_0}(\delta) = \overline{\alpha}(\delta)c_0 + \{1 - \overline{\alpha}(\delta)\}d_0$ satisfies NS($\overline{M}_0(\delta)$), where*

$$\overline{M}_0(\delta) := \Big\{ q \in \mathcal{P}(\tilde{S}_0^\infty) \mid \sup_{\substack{\mathbf{X} \in \{C,D\}^k,\ k=1,2,\ldots \\ q(\mathbf{X}C_+ \cup \mathbf{X}D_+)>0}} \frac{q(\mathbf{X}C_+)}{q(\mathbf{X}C_+ \cup \mathbf{X}D_+)} < \overline{\alpha}(\delta) \Big\}.$$

To add, consider another refinement concept, *belief-free* (e.g., Ely et al., 2005). In any equilibrium consisting of match-independent strategies, beliefs regarding the past (unobserved) behavior of a new opponent does not matter. Hence, our tolerant equilibria as well as all existing equilibria in the literature, in particular the symmetric trust-building equilibria in Greve-Okuno, are belief-free. This is a contrasting result to Heller (2017). He showed that in ordinary repeated games, only the trivial equilibria of the repetition of the one-shot Nash equilibrium may satisfy evolutionary stability or neutral stability. The difference between our result and Heller's comes from the special "imperfect monitoring" structure of our model: no information at the time of a new match and perfect monitoring during a partnership.

3.6 Stability comparison with other equilibria

There are numerous other Nash equilibria in the VSRPD. One salient class is strategy distributions consisting of "always-D"-strategies. As shown in the companion paper, Fujiwara-Greve and Okuno-Fujiwara (2019), the monomorphic d_0-distribution, which is a Nash equilibrium at any $\delta \in (0,1)$, does not satisfy External Stability nor NS($\overline{M}_0(\delta)$). We show that any distribution consisting of tolerant but always-D-strategies is also less stable than any tolerant equilibrium.

Proposition 4 *Take any Nash equilibrium p such that supp(p) $\subset \{D^k_{d_0} \mid k \in \{0, 1, \ldots, K\}\}$ for some $K < \infty$.*

(i) For any $\delta \in (0, 1)$, p does not satisfy S-NS($\mathbf{S}$).

(ii) Let $\tau(p) = \max\{k : D^k_{d_0} \in supp(p)\}$ be the maximal duration of the tolerant phase in the support. For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, p does not satisfy External Stability with respect to $M_T(\delta)$ for any $T = 0, 1, 2, \ldots, \tau$.

Proof. See Appendix.

For Proposition 4 (i), we show that there is a secret-handshake strategy (which tolerate one period longer than all existing players and cooperate if the partnership continues) that earns a higher post-entry payoff than the most tolerant $D^k_{d_0}$, so the original Nash equilibrium is not a S-NS($\mathbf{S}$). For (ii), we show that even if we restrict mutant distributions to be "dispersed", there is a bimorphic mutant distribution $q = \alpha \cdot D^{\tau+1}_{c_0} + (1 - \alpha) D^{\tau+1}_{d_0}$ such that $q \in M_T(\delta)$ for any $T = 0, 1, 2, \ldots, \tau$. Using the $D^{\tau+1}_{c_0}$-strategy as a "spring board"[26], the $D^{\tau+1}_{d_0}$-strategy outperforms all constituent strategies in *supp*(p).

Let us turn to the trust-building equilibria in Greve-Okuno, which satisfy S-NS($\mathbf{S}$) for sufficiently long initial trust-building phase.

Definition 13 *For each $T = 1, 2, \ldots,$ let the c_T-strategy be a strategy such that*

$t \leqq T$ (Trust-building phase): Play D and keep if and only if (D, D) is observed in that period;

$t \geqq T + 1$ (Cooperation phase): Play C and keep if and only if (C, C) is observed in that period.

Proposition 5 There exists $\underline{\delta}(\infty) \in (0, 1)$ such that for any $\delta \in (\underline{\delta}(\infty), 1)$, there exists $\tau(\delta)$ such that the monomorphic distribution of c_T-strategy satisfies S-NS($\mathbf{S}$) for any $T > \tau(\delta)$, but does not satisfy External Stability with respect to $M_t(\delta)$ for any $t = 0, 1, 2, \ldots, T - 1$, i.e., there exists $q \in M_t(\delta)$ and $s \in supp(q)$ such that

$$v(s; (1 - \epsilon)c_T + \epsilon \cdot q) > v(c_T; (1 - \epsilon)c_T + \epsilon \cdot q), \ \forall \epsilon \in (0, 1).$$

[26] This argument is different from the indirect invasion in van Veelen (2012). The $D^{\tau+1}_{d_0}$-mutants need to bring some $D^{\tau+1}_{c_0}$-mutants in order to beat the incumbents.

Proof. See Appendix. The bounds $\underline{\delta}(\infty)$ and $\tau(\delta)$ are constructed in Greve-Okuno.

The "intolerance" in the trust-building phase makes the monomorphic c_T-distribution vulnerable to some **bimorphic** mutant distributions containing the tolerant $D_{c_0}^T$-strategy and an earlier cooperator $D^{T-1}C_{c_0}$-strategy. Although the c_T- and $D_{c_0}^T$-strategy have the same play path when they meet among themselves, the tolerant mutant $D_{c_0}^T$-strategy does not end a partnership with the earlier cooperator $D^{T-1}C_{c_0}$-strategy, while the intolerant c_T-player does, missing the opportunity to establish a long-term cooperative relationship. Like the logic of Proposition 4 (ii), the tolerant $D_{c_0}^T$-strategy can use the earlier cooperator $D^{T-1}C_{c_0}$-strategy as the spring board to outperform the incumbent c_T-strategy. The stability of the monomorphic c_T-strategy hinged on the restriction of monomorphic mutants. (Proposition 5 implies that the c_T-equilibrium does not satisfy S-NS($M_t(\delta)$) for any $t = 0, 1, 2, \ldots, T-1$, either.)

A tolerant monomorphic trust-building equilibrium $D_{c_0}^T$ for sufficiently long T is not so stable, either. It is vulnerable to indirect invasion (van Veelen, 2012) because the intolerant c_T-strategy can emerge in the population (since they have the same play path among themselves) and then get defeated by the mutant distributions in Proposition 5.

4 Concluding Remarks

We discuss how the logic of this paper may extend to other stage games with voluntary separation, how our result can be interpreted in the network context, and future research directions.

The construction of the fundamentally asymmetric equilibrium is thanks to the Prisoner's Dilemma payoff structure (non-myopic action combination is beneficial if it is repeated) and the voluntary nature of the partnerships (cooperators can avoid exploitation). By contrast, the extension to the tolerant equilibria is due to the recursive structure of the dynamic game. With the stationary birth/death process and a fixed stage game, we can internalize the matching pool distribution within partnerships. Hence, for any stage game, it is possible to construct tolerant equilibria based on equilibria with on-path separation. However, the most striking result is for the Prisoner's Dilemma. (It is also possible to construct tolerant versions of polymorphic trust-building equilibria in the VSRPD, for example adding initial tolerant phase to an equilibrium consisting of some c_T- and c_{T+1}-strategy. However such tolerant trust-building equilibria do not exhibit as diverse behavior patterns as those based on the c_0-d_0 equilibrium, because in the former, (D, D) (as well as (C, C)) leads to mutual cooperation.)

In the context of *social games*, the VSRPD approach is a first step towards a unified analysis of network formation and within-network strategic behavior. There is a large literature of network formation (see for example, Jackson, 2008), but the strategic behavior within a network and dynamic change of the network are usually separately analyzed.[27] We showed that a huge variety of pairwise cooperative networks (between $\mathbf{X}_{c_0}$-players) and non-networking d_0-players[28] can co-exist in the society over the long horizon. This also implies that it is not guaranteed that all agents in the society end up in a (long-term) network (cf. Cho and Matsui, 2012).

There are many interesting future research directions. Although we motivated the VSRPD framework by globalization, advancement of information technology should give some information about newly matched opponents. Then an equilibrium must take into account the cost of such information and the benefit of improved punishments. Note, however, that improved information may tempt the society to aim for a personalized-punishment system to provide cooperation incentives. But personalized-punishment strategies are not belief-free, and do not provide a "safety net" property of partnership-independent equilibria like tolerant equilibria.

Another important extension is a two population model of firms and workers, to make a closed model of efficiency wage theory (e.g., Okuno, 1981[29] and Shapiro and Stiglitz, 1984). If there is an equilibrium in which a cooperative strategy (to work hard) and a myopic strategy (to shirk) co-exist among workers, it would give a further rationale to equilibrium unemployment.

Appendix

[27] An exception is Immorlica et al. (2014).

[28] The exploiters, $\mathbf{X}_{d_0}$-players, can be an intermediate type of short-term networking players.

[29] The English version is Okuno-Fujiwara (1987), but we cite the Japanese version to show that it precedes Shapiro and Stiglitz (1984).

Proof of Lemma 1. By definition,

$$\begin{aligned} & v(s;p) > v(\hat{s};p) \\ \iff & \frac{U(s;p)}{L(s;p)} > \frac{U(s;p) + \Delta U(\hat{s},s;p)}{L(s;p) + \Delta L(\hat{s},s;p)} \\ \iff & U(s;p) \cdot \Delta L(\hat{s},s;p) > L(s;p) \cdot \Delta U(\hat{s},s;p), \end{aligned} \tag{15}$$

because $L(s;p) \geqq 1$ and $L(\hat{s};p) = L(s;p) + \Delta L(\hat{s},s;p) \geqq 1$ for any $s, \hat{s}, p$.

If $\Delta L(\hat{s},s;p) > 0$, then the inequality (15) is equivalent to

$$\frac{U(s;p)}{L(s;p)} > \frac{\Delta U(\hat{s},s;p)}{\Delta L(\hat{s},s;p)}.$$

If $\Delta L(\hat{s},s;p) < 0$, then (15) is equivalent to

$$\frac{U(s;p)}{L(s;p)} < \frac{\Delta U(\hat{s},s;p)}{\Delta L(\hat{s},s;p)} = \frac{-\Delta U(\hat{s},s;p)}{-\Delta L(\hat{s},s;p)}.$$

Q.E.D.

Proof of Lemma 2.

We show the "max" part by induction on J. The "min" part is analogous.

The statement clearly holds for $J = 1$. Suppose that the claim holds for $J = n - 1$.

Take any $A_1, \ldots, A_n \in \mathbb{R}$ and any $B_1, \ldots, B_n \in \mathbb{R}_{++}$.

Let $\frac{A_i}{B_i} := \max_{j=1,2,\ldots,n} \frac{A_j}{B_j}$. We want to show that

$$\frac{\sum_{j=1}^n A_j}{\sum_{j=1}^n B_j} = \frac{A_i + \sum_{j\in\{1,2,\ldots,n\}\setminus\{i\}} A_j}{B_i + \sum_{j\in\{1,2,\ldots,n\}\setminus\{i\}} B_j} \leqq \frac{A_i}{B_i}$$

which is equivalent to

$$\frac{\sum_{j\in\{1,2,\ldots,n\}\setminus\{i\}} A_j}{\sum_{j\in\{1,2,\ldots,n\}\setminus\{i\}} B_j} \leqq \frac{A_i}{B_i}.$$

(This equivalence uses the assumption that all B_j's are positive.) Since the claim holds for $J = n-1$,

$$\frac{\sum_{j\in\{1,2,\ldots,n\}\setminus\{i\}} A_j}{\sum_{j\in\{1,2,\ldots,n\}\setminus\{i\}} B_j} \leqq \max_{j\in\{1,2,\ldots,n\}\setminus\{i\}} \{\frac{A_j}{B_j}\}.$$

By the definition,

$$\max_{j\in\{1,2,\ldots,n\}\setminus\{i\}} \{\frac{A_j}{B_j}\} \leqq \max_{j=1,\ldots,n} \frac{A_j}{B_j} = \frac{A_i}{B_i}.$$

Therefore,

$$\frac{\sum_{j\in\{1,2,\ldots,n\}\setminus\{i\}} A_j}{\sum_{j\in\{1,2,\ldots,n\}\setminus\{i\}} B_j} \leqq \frac{A_i}{B_i}.$$

Q.E.D.

Before giving the proof of Proposition 2, recall the concept of *induced* tolerant strategies, which makes payoff comparisons easy, and a lemma that gives the payoff difference between a tolerant strategy and its induced less-tolerant strategy.

Lemma 6 *For any $k = 1, 2, \ldots$, any k-period tolerant strategy $\mathbf{X}_z \in \tilde{S}_k$, its induced $(k-1)$-period tolerant strategy $\mathbf{X}'_{x_k} \in \tilde{S}_{k-1}$ and any $p \in \mathcal{P}(\tilde{S}_0^\infty)$, the total expected payoff difference is*

$$\Delta U(\mathbf{X}_z, \mathbf{X}'_{x_k}; p) \ = \delta^{2k} \sum_{\substack{\tilde{\mathbf{X}}\in\{C,D\}^k \\ (X_t,\tilde{X}_t)\neq(C,C)\ \forall t=1,\ldots,k}} [p(\tilde{\mathbf{X}}C_+)U(z,c_0) + p(\tilde{\mathbf{X}}D_+)U(z,d_0)]$$

Proof of Lemma 6. Fix an arbitrary $k \in \{1, 2, \ldots\}$. For notational simplicity, let $\hat{s} = \mathbf{X}_z$ be a k-period tolerant strategy and its induced $(k-1)$-period tolerant strategy be $s^* = \mathbf{X}'_{x_k}$. Fix an arbitrary τ-period tolerant strategy $s = \tilde{\mathbf{X}}_{\tilde{z}}$ as the partner.

Step 1: If $\tau \leqq k-1$, then $\hat{s}$ and s^* have the same payoff sequence in the partnership with s:

$$\forall \tau \leqq k-1,\ \forall\ \tilde{\mathbf{X}}_{\tilde{z}} \in \tilde{S}_\tau,\ U(\hat{s}, \tilde{\mathbf{X}}_{\tilde{z}}) - U(s^*, \tilde{\mathbf{X}}_{\tilde{z}}) = 0. \quad (16)$$

Proof of Step 1: In the first period of the match with s, both $\hat{s}$ and s^* obtain the one-shot payoff of $u(X_1, \tilde{X}_1)$. If $(X_1, \tilde{X}_1) = (C, C)$, then both $\hat{s}$ and s^* obtain the sequence of payoffs $c, c, \ldots$ as long as the partners lives.

If $(X_1, \tilde{X}_1) \neq (C, C)$, both partners survive, and $\tau > 0$, then in the second period of the match with $s = \tilde{\mathbf{X}}_{\tilde{z}}$, both $\hat{s}$ and s^* obtain the same one-shot payoff of $u(X_2, \tilde{X}_2)$. If $(X_2, \tilde{X}_2) = (C, C)$, then both $\hat{s}$ and s^* obtain the sequence of payoffs $c, c, \ldots$ as long as the partners live.

This is repeated until the τ-th period of the match as long as (C, C) is not established and both partners survive. In the $\tau + 1$ period of the match after $(X_\tau, \tilde{X}_\tau) \neq (C, C)$, s commits to $\tilde{z} \in \{c_0, d_0\}$.

Since $\hat{s}$ and s^* have not committed and have the same action plan at $\tau+1$, the payoff sequences that $\hat{s}$ and s^* obtain with s are the same. $\square$

Using the same logic, if the partner is a k-period tolerant strategy, the payoff difference of $\hat{s}$ and s^* lies only in $k+1$-th period and when (C,C) has not been established:

$$\forall \tilde{\mathbf{X}}_{\tilde{z}} \in \tilde{S}_k,\ U(\hat{s},\tilde{\mathbf{X}}_{\tilde{z}}) - U(s^*,\tilde{\mathbf{X}}_{\tilde{z}}) = \begin{cases} 0 & \text{if } (X_t,\tilde{X}_t)=(C,C),\ \exists\, t=1,2,\dots k, \\ \delta^{2k}U(z,\tilde{z}) & \text{otherwise,} \end{cases} \tag{17}$$

where z is the commitment strategy of $\hat{s}$.

Next, we show that, among k-period or more tolerant partners, it is sufficient to look at k-period tolerant partners, to compare the payoff of $\hat{s}$ and s^*.

Step 2: For any k -period action sequence $\tilde{\mathbf{X}} \in \{C,D\}^k$,

$$\tilde{\mathbf{X}}_{\tilde{z}} \in \tilde{\mathbf{X}}C_+ \Rightarrow U(\hat{s},\tilde{\mathbf{X}}_{\tilde{z}}) = U(\hat{s},\tilde{\mathbf{X}}_{c_0}),$$

$$\tilde{\mathbf{X}}_{\tilde{z}} \in \tilde{\mathbf{X}}D_+ \Rightarrow U(\hat{s},\tilde{\mathbf{X}}_{\tilde{z}}) = U(\hat{s},\tilde{\mathbf{X}}_{d_0}).$$

Proof of Step 2: Recall that $\hat{s}$ commits to one of the c_0- or the d_0-strategies in $k+1$-th period at latest. If $(X_t,\tilde{X}_t)=(C,C)$ at some $t \leqq k$, clearly $U(\hat{s},\tilde{\mathbf{X}}_{\tilde{z}}) = U(\hat{s},\tilde{\mathbf{X}}_{c_0}) = U(\hat{s},\tilde{\mathbf{X}}_{d_0})$ for any $\tilde{z}$, since $k+1$-period-plan does not matter. If $(X_t,\tilde{X}_t) \neq (C,C)$ for all $t \leqq k$, in $k+1$-th period, the continuation payoff sequence of $\hat{s}$ depends only on the $k+1$-th period action by the partner, because $\hat{s}$ plays either c_0 or d_0. Hence any $\tilde{z} \in C_+$ (resp. $\tilde{z} \in D+$) gives the same payoff as c_0 (resp. d_0) by the partner. *Q.E.D.*

Proof of Proposition 2. Fix an arbitrary $\delta \in (\underline{\delta}_{c_0d_0},1)$, an arbitrary $T=1,2,\dots$, and an arbitrary $\bar{p}_T \in P_T$. As in the one-period tolerant equilibrium example, it suffices to show the payoff equivalence of all constituent strategies:

$$\forall k=0,1,\dots,T,\ \forall \mathbf{X}_z \in supp(\bar{p}_T),\ v(\mathbf{X}_z;\bar{p}_T) = v_0(c_0;\overline{\alpha}(\delta)) = v_0(d_0;\overline{\alpha}(\delta)).$$

By mathematical induction, we show that any k-period tolerant strategy and its induced $(k-1)$-period tolerant strategy have the same payoff under $\bar{p}_T$.

Consider the c_0- and d_0-strategy, if they are in the support of the $\bar{p}_T$. The share of C-start (tolerant) strategies in the matching pool determines their average payoffs. By construction, $\bar{p}_T$ satisfies $\bar{p}_T(C_+) = \frac{\bar{p}_T\big(C_+\big)}{\bar{p}_T\big(C_+ \cup D_+\big)} = \overline{\alpha}(\delta)$. Hence

$$v(c_0; \bar{p}_T) = v(c_0; \bar{p}_T(C_+)) = v_0(c_0; \overline{\alpha}(\delta)) = v_0(d_0; \overline{\alpha}(\delta)) = v(d_0; \bar{p}_T). \tag{18}$$

Take any 1-period tolerant strategy $X_z \in supp(\bar{p}_T)$ and its induced 0-period tolerant strategy

$$x = \begin{cases} c_0 & \text{if } X = C \\ d_0 & \text{if } X = D. \end{cases}$$

By Lemmas 2 and 6 and the definition of $\bar{p}_T$,

$$\begin{aligned}
\frac{\Delta U(X_z, x; \bar{p}_T)}{\Delta L(X_z, x; \bar{p}_T)} &= \frac{\delta^2 \sum_{\substack{\tilde{X} \in \{C,D\} \\ (X,\tilde{X}) \neq (C,C)}} \bar{p}_T(\tilde{X}C_+)U(z, c_0) + \bar{p}_T(\tilde{X}D_+)U(z, d_0)}{\delta^2 \sum_{\substack{\tilde{X} \in \{C,D\} \\ (X,\tilde{X}) \neq (C,C)}} \bar{p}_T(\tilde{X}C_+)L(z, c_0) + \bar{p}_T(\tilde{X}D_+)L(z, d_0)} \\
&= \frac{\sum_{\substack{\tilde{X} \in \{C,D\} \\ (X,\tilde{X}) \neq (C,C)}} \bar{p}_T(\tilde{X}C_+ \cup \tilde{X}D_+)\{\overline{\alpha}(\delta)U(z, c_0) + (1 - \overline{\alpha}(\delta))U(z, d_0)\}}{\sum_{\substack{\tilde{X} \in \{C,D\} \\ (X,\tilde{X}) \neq (C,C)}} \bar{p}_T(\tilde{X}C_+ \cup \tilde{X}D_+)\{\overline{\alpha}(\delta)L(z, c_0) + (1 - \overline{\alpha}(\delta))L(z, d_0)\}} = v_0(z; \overline{\alpha}(\delta)).
\end{aligned}$$

Since $z, x \in \{c_0, d_0\}$,

$$\frac{\Delta U(X_z, x; \bar{p}_T)}{\Delta L(X_z, x; \bar{p}_T)} = v_0(x; \overline{\alpha}(\delta)).$$

By Lemma 1 (i), this implies that

$$v(X_z; \bar{p}_T) = v(x; \bar{p}_T).$$

Assume that for each $m = 0, 1, \ldots, k-1$, it is true that any m-period tolerant strategy $\mathbf{X}_z \in supp(\bar{p}_T)$ is payoff-equivalent to its induced $(m-1)$-period tolerant strategy $\mathbf{X}'_{x_{m-1}}$. (Then they are all payoff-equivalent to the c_0- and d_0-strategy as well.) Take any k-period tolerant strategy $\mathbf{X}_z \in supp(\bar{p}_T)$, and its induced $(k-1)$-period tolerant strategy $\mathbf{X}'_{x_k} \in \tilde{S}_{k-1}$. (If $\mathbf{X}_z \in supp(\bar{p}_T)$ does not exist, the Lemma trivially holds.) By Lemmas 2 and 6,

$$\begin{aligned}
\frac{\Delta U(\mathbf{X}_z, \mathbf{X}'_{x_k}; \bar{p}_T)}{\Delta L(\mathbf{X}_z, \mathbf{X}'_{x_k}; \bar{p}_T)} &= \frac{\sum_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^k \\ (X_t,\tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,k}} \bar{p}_T(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)U_0\Big(z; \overline{\alpha}(\delta)\Big)}{\sum_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^k \\ (X_t,\tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,k}} \bar{p}_T(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)L_0\Big(z; \overline{\alpha}(\delta)\Big)} \\
&= v_0(z; \overline{\alpha}(\delta)) = v(c_0; \bar{p}_T) = v(d_0; \bar{p}_T).
\end{aligned} \tag{19}$$

By assumption, any $(m-1)$-period tolerant strategy is payoff-equivalent to c_0 (or d_0). Hence (19) implies that

$$\frac{\Delta U(\mathbf{X}_z, \mathbf{X}'_{x_k}; \bar{p}_T)}{\Delta L(\mathbf{X}_z, \mathbf{X}'_{x_k}; \bar{p}_T)} = v(\mathbf{X}'_{x_k}; \bar{p}_T).$$

By Lemma 1 (i), this implies that $\mathbf{X}_z$ is payoff-equivalent to $\mathbf{X}'_{x_k}$, which is payoff-equivalent to the c_0- and d_0-strategy. Finally, recall Lemma 4 (Lemma 5 of GOS) which implies that payoff-equivalence guarantees that the distribution is a Nash equilibrium. This completes the proof of Proposition 2. *Q.E.D.*

Before going to the proof of Lemma 5, we give bounds to the average payoff difference between a k-period tolerant strategy and its induced m-period tolerant strategy (see Definition 11) for any $m < k$.

Lemma 7 *Fix any $k = 1, 2, \ldots$ and any $m < k$. Take any k-period tolerant strategy $\mathbf{X}_z \in \tilde{S}_k$ and its induced m-period tolerant strategy $\mathbf{X}'_{x_{m+1}} \in \tilde{S}_m$. Then, for any $p \in \mathcal{P}(\tilde{S}_0^\infty)$,*

$$\frac{\Delta U(\mathbf{X}_z, \mathbf{X}'_{x_{m+1}}; p)}{\Delta L(\mathbf{X}_z, \mathbf{X}'_{x_{m+1}}; p)} \leqq \max_{j=m+1,\ldots,k} \max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C), \forall t=1,\ldots,j}} v_0\Big(x_{j+1}; \frac{p(\tilde{\mathbf{X}}C_+)}{p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big); \tag{20}$$

$$\frac{\Delta U(\mathbf{X}_z, \mathbf{X}'_{x_{m+1}}; p)}{\Delta L(\mathbf{X}_z, \mathbf{X}'_{x_{m+1}}; p)} \geqq \min_{j=m+1,\ldots,k} \min_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C), \forall t=1,\ldots,j}} v_0\Big(x_{j+1}; \frac{p(\tilde{\mathbf{X}}C_+)}{p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big), \tag{21}$$

where

$$x_{j+1} = \begin{cases} c_0 & \text{if } X_{j+1} = C \\ d_0 & \text{if } X_{j+1} = D, \end{cases}$$

for $j = m+1, \ldots, k-1$, and $x_{k+1} = z$.

Proof of Lemma 7. Take any $m < k$ and a k-period tolerant strategy $s_k := \mathbf{X}_z \in \tilde{S}_k$. For each $j = m, m+1, \ldots, k-1$, s_k's induced j-period tolerant strategy $s_j = \mathbf{X}'_{x_{j+1}} \in \tilde{S}_j$ is defined by

$$\mathbf{X}' = (X_1, \ldots, X_j);$$
$$x_{j+1} = \begin{cases} c_0 & \text{if } X_{j+1} = C \\ d_0 & \text{if } X_{j+1} = D. \end{cases}$$

For any $p \in \mathcal{P}(\tilde{S}_0^\infty)$, the payoff difference between s_k and s_m can be decomposed as

$$\Delta U(s_k, s_m; p) = \Delta U(s_k, s_{k-1}; p) + \Delta U(s_{k-1}, s_{k-2}; p) + \ldots + \Delta U(s_{m+1}, s_m; p) = \sum_{j=m+1}^{k} \Delta U(s_j, s_{j-1}; p).$$

Similarly, the expected partnership length difference is decomposed as

$$\Delta L(s_k, s_m; p) = \Delta L(s_k, s_{k-1}; p) + \Delta L(s_{k-1}, s_{k-2}; p) + \ldots + \Delta L(s_{m+1}, s_m; p) = \sum_{j=m+1}^{k} \Delta L(s_j, s_{j-1}; p).$$

By Lemma 6, for each $j = m+1, m+2, \ldots, k$,

$$\frac{\Delta U(s_j, s_{j-1}; p)}{\Delta L(s_j, s_{j-1}; p)} = \frac{\sum_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,j}} \delta^{2j}\ p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+) U_0\Big(x_{j+1}; \frac{p(\tilde{\mathbf{X}}C_+)}{p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big)}{\sum_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,j}} \delta^{2j}\ p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+) L_0\Big(x_{j+1}; \frac{p(\tilde{\mathbf{X}}C_+)}{p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big)}.$$

(Note that $x_{k+1} = z$.) By Lemma 2, for each $j = m+1, m+2, \ldots, k$, $\frac{\Delta U(s_j, s_{j-1}; p)}{\Delta L(s_j, s_{j-1}; p)}$ is bounded as follows.

$$\frac{\Delta U(s_j, s_{j-1}; p)}{\Delta L(s_j, s_{j-1}; p)} \leqq \max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,j}} v_0\Big(x_{j+1}; \frac{p(\tilde{\mathbf{X}}C_+)}{p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big). \tag{22}$$

Furthermore, by Lemma 2 again,

$$\frac{\Delta U(s_k, s_m; p)}{\Delta L(s_k, s_m; p)} = \frac{\sum_{j=m+1}^{k} \Delta U(s_j, s_{j-1}; p)}{\sum_{j=m+1}^{k} \Delta U(s_j, s_{j-1}; p)} \leqq \max_{j=m+1,\ldots,k} \frac{\Delta U(s_j, s_{j-1}; p)}{\Delta L(s_j, s_{j-1}; p)}. \tag{23}$$

(22) and (23) imply that

$$\frac{\Delta U(s_k, s_m; p)}{\Delta L(s_k, s_m; p)} \leqq \max_{j=m+1,\ldots,k} \frac{\Delta U(s_j, s_{j-1}; p)}{\Delta L(s_j, s_{j-1}; p)} \leqq \max_{j=m+1,\ldots,k} \max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,j}} v_0\Big(x_{j+1}; \frac{p(\tilde{\mathbf{X}}C_+)}{p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big).$$

This is (20). Derivation of (21) is analogous. *Q.E.D.*

Proof of Lemma 5. Fix an arbitrary $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, an arbitrary $T = 0, 1, 2, \ldots$, an arbitrary $q \in M_T(\delta)$, and any $s_T \in \tilde{S}_T$. (Since $\bar{p}_T$ is full-support, $s_T \in supp(\bar{p}_T)$.) Let

$$\hat{\alpha} := \sup_{\substack{\mathbf{X} \in \{C,D\}^{T+k},\ k=1,2,\ldots \\ q\left(\mathbf{X}C_+ \cup \mathbf{X}D_+\right) > 0}} \frac{q\big(\mathbf{X}C_+\big)}{q\big(\mathbf{X}C_+ \cup \mathbf{X}D_+\big)} (< \overline{\alpha}(\delta)).$$

Step 1: There exists $\overline{\epsilon}(s_T) \in (0,1)$ such that for any $\epsilon \in (0, \overline{\epsilon}(s_T))$,

$$v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) > \max\{v_0(c_0; \hat{\alpha}), v_0(d_0; \hat{\alpha})\}.$$

Proof of Step 1: This is straightforward from

$$\lim_{\epsilon \to 0} v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) = v_0(c_0; \overline{\alpha}(\delta)) = v_0(d_0; \overline{\alpha}(\delta))$$

and the fact that $v_0(c_0;\alpha)$ and $v_0(d_0;\alpha)$ are increasing in α, so that $\max\{v_0(c_0;\hat{\alpha}), v_0(d_0;\hat{\alpha})\} < v_0(c_0;\overline{\alpha}(\delta)) = v_0(d_0;\overline{\alpha}(\delta))$. □

Without loss of generality, assume that $S(s_T) \cap supp(q) \neq \emptyset$ (if not, the statement (13) is trivial) and take any $s' = \mathbf{X}_z \in S(s_T) \cap supp(q)$ (thus $|\mathbf{X}| = T + k$ for some $k \geqq 1$ and $z \in \{c_0, d_0\}$). Since $S(s_T) \cap supp(q) \neq \emptyset$, $q(\mathbf{X}C_+ \cup \mathbf{X}D_+) > 0$. Below, write $p^{PE}(\epsilon) = (1-\epsilon)\bar{p}_T + \epsilon \cdot q$ for any $\epsilon \in (0,1)$.

Step 2:

$$\forall \epsilon \in (0,1),\ \frac{\Delta U(s', s_T; p^{PE}(\epsilon))}{\Delta L(s', s_T; p^{PE}(\epsilon))} < \max\{v_0(c_0;\hat{\alpha}), v_0(d_0;\hat{\alpha})\}. \tag{24}$$

Proof of Step 2: The bound (20) in Lemma 7 implies that

$$\frac{\Delta U(s', s_T; p^{PE}(\epsilon))}{\Delta L(s', s_T; p^{PE}(\epsilon))} \leqq \max_{j=T+1,\dots,T+k} \max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\dots,j}} v_0\Big(x_{j+1}; \frac{p^{PE}(\tilde{\mathbf{X}}C_+)}{p^{PE}(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big),$$

where $x_{j+1} = c_0$ (resp. d_0) if $X_{j+1} = C$ (resp. D) for each $j = T+1, \dots, T+k-1$, and $x_{T+k+1} = z$. Since $\tilde{\mathbf{X}}C_+$ and $\tilde{\mathbf{X}}D_+$ are subsets of $T+1$-period and longer tolerant strategies, $p^{PE}(\tilde{\mathbf{X}}C_+) = \epsilon \cdot q(\tilde{\mathbf{X}}C_+)$ and $p^{PE}(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+) = \epsilon \cdot q(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)$.[30] Hence the above inequality is equivalent to

$$\frac{\Delta U(s', s_T; p^{PE}(\epsilon))}{\Delta L(s', s_T; p^{PE}(\epsilon))} \leqq \max_{j=T+1,\dots,T+k} \max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\dots,j}} v_0\Big(x_{j+1}; \frac{q(\tilde{\mathbf{X}}C_+)}{q(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big).$$

Thus the upper bound is independent of ϵ.

Since $v_0(c_0;\alpha)$ and $v_0(d_0;\alpha)$ are increasing in α, and by the assumption $q \in M_T(\delta)$, for each $j = T+1, \dots, T+k$,

$$\max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\dots,j}} v_0\Big(x_{j+1}; \frac{q(\tilde{\mathbf{X}}C_+)}{q(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big) \leqq v_0(x_{j+1}; \hat{\alpha}) \leqq \max\{v_0(c_0;\hat{\alpha}), v_0(d_0;\hat{\alpha})\},$$

[30] This is not always true if the equilibrium $\bar{p}_T$ is not full support. For example, if some $k \leqq T$-period tolerant strategies are absent in $\bar{p}_T$, then the class of $\tilde{\mathbf{X}}C_+$ can contain both equilibrium strategies and mutant strategies.

i.e., (24) holds. $\square$

Finally, note that $\tilde{S}_T$ is a finite set. Hence there exists

$$\bar{\epsilon}_T = \min_{s_T \in \tilde{S}_T} \bar{\epsilon}(s_T),$$

and by Steps 1 and 2, for any $\epsilon \in (0, \bar{\epsilon}_T)$, any $s_T \in \tilde{S}_T$, and any $s' \in S(s_T) \cap supp(q)$,

$$\frac{\Delta U(s', s_T; p^{PE}(\epsilon))}{\Delta L(s', s_T; p^{PE}(\epsilon))} \leqq \max\{v_0(c_0; \hat{\alpha}), v_0(d_0; \hat{\alpha})\} < v(s_T; p^{PE}(\epsilon)).$$

By Lemma 1 (i),

$$v(s'; p^{PE}(\epsilon)) < v(s_T; p^{PE}(\epsilon)). \qquad Q.E.D.$$

Proof of Proposition 3. Fix any $T \in \{0, 1, 2, \ldots\}$, any $\bar{p}_T \in P_T^\circ$ and any $m \in \{T+1, T+2, \ldots\}$.

Step 1: There exists (in fact many) $q \in M_T(\delta) \cap int(\mathcal{P}(\tilde{S}_0^m))$ such that

$$\forall k = 0, 1, \ldots, T-1, \ \forall \ s_k \in \tilde{S}_k, \ q(s_k) = \bar{p}_T(s_k), \tag{25}$$

$$\forall s_T \in \tilde{S}_T, \ q(s_T) + \sum_{s' \in S(s_T)} q(s') = \bar{p}_T(s_T). \tag{26}$$

Proof of Step 1: Recall that $M_T(\delta)$ does not restrict $q(s)$ for any $s \in \tilde{S}_k$ for any $k = 0, 1, \ldots, T$, and only the relative ratio among $s \in \tilde{S}_k$ for $k = T+1, T+2, \ldots$. Hence there exists q which satisfies the above equalities in $M_T(\delta)$. Moreover, such q can be a "full-support" $q \in int(\mathcal{P}(\tilde{S}_0^m))$ because $M_T(\delta)$ does not require that any strategy has a 0 probability, and since $\bar{p}_T$ is full-support, $\bar{p}_T(s_k) > 0$ for any $k = 0, 1, \ldots, T$. $\square$

Step 2: The above q satisfies (12). That is, there exists $\bar{\epsilon}_T \in (0, 1)$ such that for any $\epsilon \in (0, \bar{\epsilon}_T)$,

$$\sum_{s \in supp(\bar{p}_T)} \bar{p}_T(s) v(s; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) > \sum_{s' \in supp(q)} q(s') v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q). \tag{12}$$

Proof of Step 2: Since $q \in M_T(\delta)$, External Stability Lemma 5 implies that there exists $\bar{\epsilon}_T \in (0, 1)$ such that for any $\epsilon \in (0, \bar{\epsilon}_T)$,

$$\forall s_T \in \tilde{S}_T, \ \forall s' \in S(s_T) \cap supp(q), \quad v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) > v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q).$$

Therefore, for each $s_T \in \tilde{S}_T$,

$$\begin{aligned}
&\sum_{s' \in S(s_T) \cap supp(q)} q(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) + q(s_T)v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
< &\sum_{s' \in S(s_T) \cap supp(q)} q(s')v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) + q(s_T)v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
= &[q(s_T) + \sum_{s' \in S(s_T) \cap supp(q)} q(s')] \cdot v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
= &\bar{p}_T(s_T) \cdot v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \qquad \text{from (26).}
\end{aligned}$$

Summing up both sides for all s_T's,

$$\begin{aligned}
&\sum_{s_T \in \tilde{S}_T} q(s_T)v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) + \sum_{s_T \in \tilde{S}_T} \sum_{s' \in S(s_T) \cap supp(q)} q(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
< &\sum_{s_T \in \tilde{S}_T} \bar{p}_T(s_T) \cdot v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q).
\end{aligned}$$

From (25), for any $k = 0, 1, \ldots, T-1$ and $s_k \in \tilde{S}_k$,

$$q(s_k)v(s_k; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) = \bar{p}_T(s_k)v(s_k; (1-\epsilon)\bar{p}_T + \epsilon \cdot q).$$

Adding all these gives

$$\begin{aligned}
&\sum_{k=0}^{T-1} \sum_{s_k \in \tilde{S}_k} q(s_k)v(s_k; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
&+ \sum_{s_T \in \tilde{S}_T} q(s_T)v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) + \sum_{s_T \in \tilde{S}_T} \sum_{s' \in S(s_T) \cap supp(q)} q(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
< &\sum_{k=0}^{T-1} \sum_{s_k \in \tilde{S}_k} \bar{p}_T(s_k)v(s_k; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) \\
&+ \sum_{s_T \in \tilde{S}_T} \bar{p}_T(s_T) \cdot v(s_T; (1-\epsilon)\bar{p}_T + \epsilon \cdot q).
\end{aligned}$$

Since $supp(q) = \tilde{S}_0^m$ and $supp(\bar{p}_T) = \tilde{S}_0^T$, the LHS is $\sum_{s' \in supp(q)} q(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q)$ and the RHS is $\sum_{s \in supp(\bar{p}_T)} \bar{p}_T(s)v(s; (1-\epsilon)\bar{p}_T + \epsilon \cdot q)$. □

Step 3: There exists a (sufficiently small) real number $r > 0$ such that for any $q' \in int(\mathcal{P}(\tilde{S}_0^m))$ such that $|q - q'| < r$, $q' \in M_T(\delta)$ and

$$\sum_{s \in supp(\bar{p}_T)} \bar{p}_T(s)v(s; (1-\epsilon)\bar{p}_T + \epsilon \cdot q') > \sum_{s' \in supp(q)} q'(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q'). \tag{27}$$

Proof of Step 3: Since the condition of $M_T(\delta)$ is a strict inequality, there exists $r_1 > 0$ such that for any $q' \in int(\mathcal{P}(\tilde{S}_0^m))$ such that $|q - q'| < r_1$,

$$\sup_{\substack{\mathbf{X} \in \{C,D\}^{T+k},\ k=1,2,\ldots \\ q'\left(\mathbf{X}C_+ \cup \mathbf{X}D_+\right)>0}} \frac{q'\left(\mathbf{X}C_+\right)}{q'\left(\mathbf{X}C_+ \cup \mathbf{X}D_+\right)} < \overline{\alpha}(\delta)$$

holds, i.e., $q' \in M_T(\delta)$.

Since (12) is a strict inequality, there exists another $r_2 > 0$ such that for any $q'' \in int(\mathcal{P}(\tilde{S}_0^m))$ such that $|q - q''| < r_2$,

$$\begin{aligned}
&\Big| \sum_{s' \in supp(q)} q(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) - \sum_{s' \in supp(q)} q''(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q'')\Big| \\
&< \sum_{s \in supp(\bar{p}_T)} \bar{p}_T(s)v(s; (1-\epsilon)\bar{p}_T + \epsilon \cdot q) - \sum_{s' \in supp(q)} q(s')v(s'; (1-\epsilon)\bar{p}_T + \epsilon \cdot q).
\end{aligned}$$

Take $r = \min\{r_1, r_2\}$ and we have (27). □

Step 3 shows that there is an open ball around each q in Step 1 with a non-empty interior which makes $\bar{p}_T$ mean stable. *Q.E.D.*

Proof of Proposition 4. (i) Fix an arbitrary polymorphic Nash equilibrium p such that $supp(p) \subset \{D_{d_0}^k \mid k \in \{0, 1, \ldots, K\}\}$. We show that p is not a S-NS($\mathbf{S}$) for any $\delta \in (0, 1)$. Consider the secret-handshake strategy $D_{c_0}^{\tau+1}$ against the maximal tolerant strategy in the support of p. The payoff sequences of the strategies in p, $D_{c_0}^{\tau+1}$, and $D_{d_0}^{\tau+1}$ are shown in Table 4.

Let $p^{PE} = (1-\epsilon)p + \epsilon \cdot D_{c_0}^{\tau+1}$. From Table 4,

$$\frac{\Delta U(D_{c_0}^{\tau+1}, D_{d_0}^{\tau}; p^{PE})}{\Delta L(D_{c_0}^{\tau+1}, D_{d_0}^{\tau}; p^{PE})} = c > d = v(D_{d_0}^{\tau}; p^{PE}).$$

By Lemma 1 (i),

$$\forall \epsilon \in (0, 1), v(D_{c_0}^{\tau+1}; p^{PE}) > v(D_{d_0}^{\tau}; p^{PE}).$$

Therefore p is not a S-NS($\mathbf{S}$) (although d_0, D_{d_0} etc. may not exist in p, $D_{d_0}^{\tau}$ exists).

(ii) Recall that the maximal duration of the tolerant phase among the equilibrium strategies is denoted by $\tau = \max\{k : D_{d_0}^k \in supp(p)\}$. Then $\tau \leqq K < \infty$. We show that p does not satisfy External

you \ partner	d_0	D_{d_0}	$\cdots$	$D^{\tau}_{d_0}$	$D^{\tau+1}_{c_0}$	$D^{\tau+1}_{d_0}$
d_0	d	d	$\cdots$	d	d	d
D_{d_0}	d	d,d	$\cdots$	d,d	d,d	d,d
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$D^{\tau}_{d_0}$	d	d,d	$\cdots$	$\overbrace{d,\cdots,d}^{\tau+1 \text{ times}}$	$\overbrace{d,\cdots,d}^{\tau+1 \text{ times}}$	$\overbrace{d,\cdots,d}^{\tau+1 \text{ times}}$
$D^{\tau+1}_{c_0}$	d	d,d	$\cdots$	$\overbrace{d,\cdots,d}^{\tau+1 \text{ times}}$	$\overbrace{d,\cdots,d}^{\tau+1 \text{ times}},c,c\ldots$	$\overbrace{d,\cdots,d}^{\tau+1 \text{ times}},\ell$
$D^{\tau+1}_{d_0}$	d	d,d	$\cdots$	$\overbrace{d,\cdots,d}^{\tau+1 \text{ times}}$	$\overbrace{d,\cdots,d}^{\tau+1 \text{ times}},g$	$\overbrace{d,\cdots,d}^{\tau+1 \text{ times}},d$

Table 4: Payoff sequence comparison for D-always equilibrium

Stability with respect to $M_T(\delta)$ for any $T = 0, 1, 2, \ldots, \tau$. Consider a mutant strategy distribution $q = \alpha \cdot D^{\tau+1}_{c_0} + (1-\alpha)D^{\tau+1}_{d_0}$ such that $\alpha \in (0, \overline{\alpha}(\delta))$. For any $T = 0, 1, 2, \ldots, \tau$, $q \in M_T(\delta)$. Let $p^{PE} = (1-\epsilon)p + \epsilon \cdot q$. We show that

$$\forall s \in supp(p),\ \forall \epsilon \in (0,1), v(s; p^{PE}) \leqq v(D^{\tau+1}_{d_0}; p^{PE}).$$

For any equilibrium strategy $D^k_{d_0} \in supp(p)$, the mutant strategy $D^{\tau+1}_{d_0}$ is more tolerant and hence $\Delta L(D^{\tau+1}_{d_0}, D^k_{d_0}; p^{PE}) > 0$. In view of Lemma 1, it suffices to show that

$$\frac{\Delta U(D^{\tau+1}_{d_0}, D^k_{d_0}; p^{PE})}{\Delta L(D^{\tau+1}_{d_0}, D^k_{d_0}; p^{PE})} \geqq v(D^k_{d_0}; p^{PE}) = d. \tag{28}$$

Consider $k = 0$ as an example. From Table 4,

$$\begin{aligned}
&\Delta U(D^{\tau+1}_{d_0}, d_0; p^{PE}) \\
=&(1-\epsilon)[p(D_{d_0})\delta^2 d + p(D^2_{d_0})(\delta^2 d + \delta^4 d) + \cdots + p(D^{\tau}_{d_0})(\delta^2 d + \cdots + \delta^{2\tau} d)] \\
&+ \epsilon[\delta^{2\tau} d + \delta^{2(\tau+1)}(\alpha g + (1-\alpha)d)] \\
=& (1-\epsilon)p(D^m_{d_0}; 1 \leqq m \leqq \tau)\delta^2 d + (1-\epsilon)p(D^m_{d_0}; 2 \leqq m \leqq \tau)\delta^4 d + \cdots + [(1-\epsilon)p(D^{\tau}_{d_0}) + \epsilon]\delta^{2\tau} d \\
&+ \epsilon\delta^{2(\tau+1)}[\alpha g + (1-\alpha)d].
\end{aligned}$$

you \ partner	c_T	$D^T_{c_0}$	$D^{T-1}C_{c_0}$
c_T	$\overbrace{d,\cdots,d}^{T\text{ times}},c,\ldots$	$\overbrace{d,\cdots,d}^{T\text{ times}},c,\ldots$	$\overbrace{d,\cdots,d}^{T-1\text{ times}},g$
$D^T_{c_0}$	$\overbrace{d,\cdots,d}^{T\text{ times}},c,\ldots$	$\overbrace{d,\cdots,d}^{T\text{ times}},c,\ldots$	$\overbrace{d,\cdots,d}^{T-1\text{ times}},g,c,\ldots$

Table 5: Payoff sequences of the intolerant c_T and tolerant $D^T_{c_0}$

Analogously, for an arbitrary k,

$$\Delta U(D^{\tau+1}_{d_0}, D^k_{d_0}; p^{PE}) = \sum_{n=k+1}^{\tau} p(D^m_{d_0}; n \leqq m \leqq \tau) \cdot \delta^{2n} d + \epsilon\delta^{2\tau} d + \epsilon\delta^{2(\tau+1)}[\alpha g + (1-\alpha)d];$$

$$\Delta L(D^{\tau+1}_{d_0}, D^k_{d_0}; p^{PE}) = \sum_{n=k+1}^{\tau} p(D^m_{d_0}; n \leqq m \leqq \tau) \cdot \delta^{2n} + \epsilon\delta^{2\tau} + \epsilon\delta^{2(\tau+1)}.$$

By Lemma 2, for any $\epsilon \in (0,1)$ and any $\alpha > 0$, we have

$$\frac{\Delta U(D^{\tau+1}_{d_0}, D^k_{d_0}; p^{PE})}{\Delta L(D^{\tau+1}_{d_0}, D^k_{d_0}; p^{PE})} \geqq \min\{d, v(d_0;\alpha)\} \geqq d. \tag{28}$$

This completes the proof of (ii). *Q.E.D.*

Proof of Proposition 5. The fact that a c_T-distribution satisfies S-NS(**S**) for $T \geqq \tau(\delta)$ is proved in Greve-Okuno. Consider $q = \beta D^{T-1}C_{c_0} + (1-\beta)D^T_{c_0}$. Take an arbitrary $\epsilon \in (0,1)$ and let $p^{PE} = (1-\epsilon)c_T + \epsilon \cdot q$. The payoff difference between the (intolerant) trust-building c_T-strategy and the tolerant $D^T_{c_0}$-strategy is when they meet the earlier cooperator $D^{T-1}C_{c_0}$-strategy. In $t = T$, the c_T-player receives g but ends the partnership, while the tolerant $D^T_{c_0}$-strategy keeps the partnership to establish a long-term cooperative partnership from $T+1$-the period on. See Table 5. Hence

$$\frac{\Delta U(D^T_{c_0}, c_T; p^{PE})}{\Delta L(D^T_{c_0}, c_T; p^{PE})} = c.$$

Note that this does not depend on β or ϵ. The post-entry average payoff of c_T is

$$v(c_T; p^{PE}) = \frac{(1 + \cdots + \delta^{2(T-2)})d + \delta^{2(T-1)}[\{(1-\epsilon) + \epsilon(1-\beta)\}(d + \frac{c}{1-\delta^2}) + \epsilon \cdot \beta g]}{\{(1-\epsilon) + \epsilon(1-\beta)\}\frac{1}{1-\delta^2} + \epsilon \cdot \beta \frac{1-\delta^{2T}}{1-\delta^2}}.$$

When $\beta = 0$, $v(c_T; p^{PE}) = v(c_T, c_T) = (1-\delta^{2T})d + \delta^{2T}c < c$. Hence there exists $\overline{\beta} \in (0,1)$ such that for any $\beta \in (0, \overline{\beta})$ and any $\epsilon \in (0,1)$,

$$\frac{\Delta U(D^T_{c_0}, c_T; p^{PE})}{\Delta L(D^T_{c_0}, c_T; p^{PE})} = c > v(c_T; p^{PE}).$$

Then Lemma 1 (i) implies that

$$\forall\beta \in (0, \overline{\beta}),\ \forall\epsilon \in (0,1),\ v(D^T_{c_0}; p^{PE}) > v(c_T; p^{PE}),$$

violating External Stability. Finally, there exists $\beta \in (0, \min\{\overline{\beta}, \overline{\alpha}(\delta)\})$ so that $q \in M_t(\delta)$ for any $t = 0, 1, 2, \ldots, T-1$. *Q.E.D.*

References

Bendor, J., Swistak, P., 1997. The Evolutionary Stability of Cooperation. Am. Pol. Sci. Rev. 91 (2), 290–307.

Carmichael, L. and MacLeod, B. 1997. Gift Giving and the Evolution of Cooperation. *International Economic Review* **38** 485–509.

Cho, I. and Matsui, A. 2012. A Dynamic Foundation of the Rawlsian Maximin Criterion. *Dynamic Games and Applications* **2** 51–70.

van Damme, E. 1987. *Stability and Perfection of Nash Equilibria.* Springer, New York.

Datta, S. 1996. Building Trust. Manuscript. London School of Economics.

Duffie, D. and Sun, Y. 2012. The Exact Law of Large Numbers for Independent Random Matching. *Journal of Economic Theory* **147** 1105–1139.

Duffie, D., Quao, L. and Sun, Y. (2018). Dynamic Directed Random Matching. *Journal of Economic Theory* **174** 124–183.

Eeckhout, J. 2006. Minorities and Endogenous Segregation. *Review of Economic Studies* **73** 31–53.

Ely, J., Hörner, J. and Olszewski, W. 2005. Belief-free Equilibria in Repeated Games. *Econometrica* **73**(2) 377–415.

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

Fudenberg, D. and Maskin, E. 1986. The Folk Theorem in Repeated Games with Discounting or with Incomplete Information. *Econometrica* **54** 533–554.

Fujiwara-Greve, T. and Okuno-Fujiwara, M. 2009. Voluntarily Separable Repeated Prisoner's Dilemma. *Review of Economic Studies* **76** 993–1021.

Fujiwara-Greve, T. and Okuno-Fujiwara, M. 2012. Behavioral Diversity in Voluntarily Repeated Prisoner's Dilemma. Available at SSRN http://dx.doi.org/10.2139/ssrn.2005115.

Fujiwara-Greve, T. and Okuno-Fujiwara, M. 2016. Diverse Behavior Patterns in a Symmetric Society with Voluntary Partnerships. Available at SSRN http://dx.doi.org/10.2139/ssrn.2343119.

Fujiwara-Greve, T. and Okuno-Fujiwara, M. 2017. Long-term Cooperation and Diverse Behavior Patterns under Voluntary Partnerships. Keio University and University of Tokyo.

Fujiwara-Greve, T. and Okuno-Fujiwara, M. 2019. On Evolutionary Stability of the Fundamentally Asymmetric Equilibrium. Mimeo. Keio University.

Fujiwara-Greve, T.、Okuno-Fujiwara, M. and Suzuki, N. 2015. Efficiency may Improve when Defectors Exist. *Economic Theory* **60** 423–460.

Garcìa, J. and van Veelen, M. 2016. In and Out of Equilibrium I: Evolution of Strategies in Repeated Games with Discounting. *Journal of Economic Theory* **161** 161–189.

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

Heller, Y. 2017. Instability of Belief-free Equilibria. *Journal of Economic Theory* **168** 261–286.

Honhon, D. and Hyndman, K. 2017. Flexibility and Reputation in Repeated Prisoner's Dilemma Games. Manuscript. University of Texas at Dallas.

Immorlica, N., Lucier, B. and Rogers, B. 2014. Cooperation in Anonymous Dynamic Social Networks. Manuscript. Microsoft Research and Washington University in St. Louis.

Izquierdo, L., Izquierdo, S. and van Veelen, M. 2018. Repeated games with the option to leave. Mimeo. University of Amsterdam.

Izquierdo, L., Izquierdo, S. and Vega-Redondo, F. 2014. Leave and Let Leave: A Sufficient Condition to Explain the Evolutionary Emergence of Cooperation. *Journal of Economic Dynamics and Control* **46** 91–113.

Jackson, M. 2008. *Social and Economic Networks.* Princeton University Press, New Jersey, USA.

Kandori, M. 1992. Social Norms and Community Enforcement. *Review of Economic Studies* **59** 63–80.

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

Kranton, R. 1996. The Formation of Cooperative Relationships. *Journal of Law, Economics & Organization* **12** 214–233.

Lee, Natalie (2018) "An Experiment: Voluntary Separation in Indefinitely Repeated Prisoner's Dilemma Game", SSRN 327669.

Lei, V., Vesely, F. and Yang, C-L. 2018. Volunary Separation as a Disciplinary Device for Long-Term Cooperation: Reconciling Theory with Evidence. Manuscript. University of Wisconsin-Milwaukee.

McAdams, D. 2011. Performance and Turnover in a Stochastic Partnership. *American Economic Journal: Microeconomics* **3** 107–142.

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

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

Maynard Smith, J. and Price, G. (1973). The Logic of Animal Conflict. *Nature* **246** 15–18.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, USA.

Okuno, M. 1981. Labor Incentives and the 'Japanese' Labor Market. *Contemporary Economics (Kikan Gendai Keizai)* **Winter** 150–162. (In Japanese.)

Okuno-Fujiwara, M. 1987. Monitoring Cost, Agency Relationship, and Equilibrium Modes of Labor Contract. *Japanese and International Economies* **1** 147–167.

Okuno-Fujiwara, M., Nishimura, N., Suzuki, N. and Fujiwara-Greve, T. (2019) Voluntary Partnerships, Tolerance and Coordination: An Experimental Study. In preparation.

Rob, R. and Yang, H. 2010. Long-term Relationships as Safeguards. *Economic Theory* **43** 143–166.

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

Samuelson, L. 1997. *Evolutionary Games and Equilibrium Selection.* MIT Press, Cambridge, MA.

Sandholm, W. 2010. *Population Games and Evolutionary Dynamics.* MIT Press, Cambridge, MA.

Schumacher, H. 2013. Imitating Cooperation and the Formation of Long-term Relationships. *Journal of Economic Theory* **148** 409–417.

Selten, R. 1975. Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. *International Journal of Game Theory* **4** 25-55.

Selten, R. 1983. Evolutionary Stability in Extensive Two-Person Games. *Mathematical Social Sciences* **5** 269–363.

Shapiro, C. and Stiglitz, J. 1984. Equilibrium Unemployment as a Worker Discipline Device. *American Economic Review* **74** 433–444.

Sun, Y. 2006. The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks. *Journal of Economic Theory* **126** 31–69.

Swinkels, J. 1992. Evolutionary Stability with Equilibrium Entrants. *Journal of Economic Theory* **57** 306–332.

van Veelen, M. 2012. Robustness against Indirect Invasions. *Games and Economic Behavior* **74** 382–393.

Vesly, F. and Yang, C. 2012. On Optimal Social Convention in Voluntary Continuation Prisoner's Dilemma Games. SSRN No. 2179063.

Watson, J. 2002. Starting Small and Commitment. *Games and Economic Behavior* **38** 176–199.