On Evolutionary Stability of the Fundamentally Asymmetric Equilibrium

Takako Fujiwara-Greve*
Dept. of Economics
Keio University

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

March 10, 2019. (Preliminary, comments welcome.)
For the most updated version, click here.

Abstract: Fujiwara-Greve et al. (2015) analyzed the *fundamentally asymmetric equilibrium* in the Voluntarily Separable Repeated Prisoner's Dilemma, where conditional cooperators and myopic defectors co-exist. It shows a new incentive mechanism when one can defect and run away without information flow to future partners: the existence of myopic defectors makes assortative matchings among cooperators attractive. A problem of this equilibrium is that it is not neutrally stable in the sense of Fujiwara-Greve and Okuno-Fujiwara (2009). We show that it satisfies a more standard neutral stability, based on the mean payoff of strategy distributions, when mutants do not coordinate heavily on cooperative strategies. (100 words)

JEL code: C73.

Keywords: neutral stability, cooperation, voluntary partnerships, repeated Prisoner's Dilemma.

*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 current globalized economy, it is more natural to consider voluntarily played long-run games than to look at the ordinary repeated games with the same opponents over time. One of the important frameworks to analyze the voluntary dynamic games is the Voluntarily Separable Repeated Prisoner's Dilemma (VSRPD henceforth) of Fujiwara-Greve and Okuno-Fujiwara (2009). In this model, players are randomly matched into pairs to play a Prisoner's Dilemma, and, at the end of each period, partners can choose simultaneously whether to stay with the same parter or not. A partnership dissolves if at least one player chooses to leave. Actions are only observable within a pair. Hence it is possible to defect and run away every period. This model describes double moral hazard problems in a large mobile society.

Early studies of the VSRPD and closely related models (e.g., Datta, 1996, Ghosh and Ray, 1996, Kranton, 1996, Carmichael and Macleod, 1997, Eeckhout, 2006, Fujiwara-Greve and Okuno-Fujiwara, 2009, and Rob and Yang, 2010) advocated initial trust-building/gradual cooperation to deter moral hazard. The initial trust-building periods where no one cooperates serve as the punishment if one defects and runs away when her/his partner starts cooperating. However, this requires that all players in the society coordinate on the sufficient number of the trust-building periods.

Fujiwara-Greve et al. (2015) analyzed the *fundamentally asymmetric equilibrium*, which does not require such a coordination in the society. In this equilibrium, conditional cooperators and myopic defectors co-exist.[1] Its logic of how cooperation can be sustained in a mobile society is different from the trust-building idea: Because of the existence of defectors in the matching pool, cooperators want to stay together and hence they cooperate. It also lays a foundation to a huge class of "tolerant" equilibria (Fujiwara-Greve and Okuno-Fujiwara, 2019) and, in a subclass of Prisoner's Dilemma, it is more efficient than any Nash equilibrium consisting of trust-building strategies.

A drawback of the fundamentally asymmetric equilibrium is that it does not satisfy the (strong) neutral stability defined in Fujiwara-Greve and Okuno-Fujiwara (2009). This stability requires that **each** strategy in the support of an equilibrium must not be outperformed by any pure-strategy adopted

[1]See also a dynamic analysis of this equilibrium by Izquierdo et al. (2014), which is inspired by our earlier work, Fujiwara-Greve and Okuno-Fujiwara (2012), that introduced the fundamentally asymmetric equilibrium.

by sufficiently small measures of mutants. However, the myopic defector strategy in the fundamentally asymmetric equilibrium is inferior to the one-period trust-building strategy, which plays D and stay in the Prisoner's Dilemma in the first period of any match, and becomes the conditional cooperator from the second period on. Such mutants conduct "secret-handshake" (Robson, 1990 and Matsui, 1991) and outperform the myopic defectors. Instead of focusing on the (strong) neutral stability, Fujiwara-Greve et al. (2015) restricted the attention to the subclass of Prisoner's Dilemma in which the fundamentally asymmetric equilibrium is more efficient than trust-building equilibria. In this class, the fundamentally asymmetric equilibrium is robust against equilibrium entrants within the trust-building strategies.

This paper considers any Prisoner's Dilemma, enlarges the set of potential mutant strategies, and shows that the fundamentally asymmetric equilibrium satisfies a more standard neutral stability[2], in the sense that the mean payoff of the equilibrium strategies is not worse than the mean payoff of the mutants, when the set of potentially feasible mutant strategy distributions is appropriately restricted. The new stability concept is the "matching-pool distribution" version of the *Izquierdo & van Veelen Neutrally Stable Distribution* in Izquierdo et al. (2018), when any mutant distribution is possible.

The paper is organized as follows. In Section 2, the model and the fundamentally asymmetric equilibrium are presented. In Section 3, we introduce the new neutral stability concept and show that fundamentally asymmetric equilibrium satisfies it under "dispersed" mutant distributions. Section 4 concludes the paper with the remark that the restriction of potential mutant distributions is not serious because the monomorphic Nash equilibrium consisting of the myopic defectors is not neutrally stable with respect to the same set of potential mutant strategy distributions.

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 (Sandholm, 2010) over a discrete time horizon. The population is homogeneous and of size 1.

[2]Since the component game is an extensive-form game, there is no Evolutionarily Stable Strategy (Maynard Smith and Price, 1973 and Maynard Smith, 1982), which requires that the equilibrium strategy distribution performs strictly better than mutants. See for example, Selten (1983).

	C	D
C	c, c	ℓ, g
D	g, ℓ	d, d

Table 1: Prisoner's Dilemma: $g > c > d > \ell$ and $2c \geqq g + \ell$.

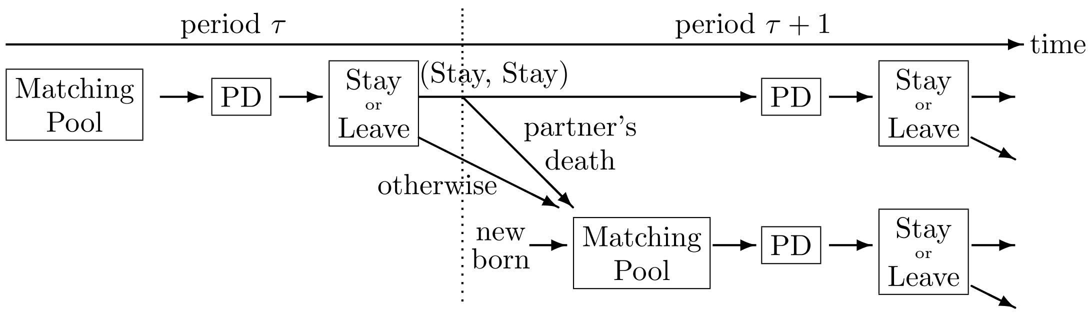

Figure 1: Timeline of the VSRPD

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.[3] Newly matched players do not have information regarding one another's past actions. Matched partners play the symmetric Prisoner's Dilemma (Table 1), where their choices are 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 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 time line is illustrated in Figure 1.

The one-shot payoffs in the Prisoner's Dilemma are in Table 1, 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.

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

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, 2006) 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[4], 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 $\tau = 1, 2, 3 \ldots$ 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, meaning that the current match is a new partnership. For each $\tau = 2, 3, \ldots$, 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

[4]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.

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 C/D and stay/leave decision rules $s = (x_{s,\tau}, y_{s,\tau})_{\tau=1}^{\infty}$ such that for all $\tau = 1, 2, \ldots$,

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

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

The set of pure private strategies is denoted by $\mathcal{S}$. Next, to define match-independent strategies, denote by $t = 1, 2, \ldots$ the period within the same partnership. If a partnership dissolves, the next partnership starts at $t = 1$.

Definition 2 For each period $t = 2, 3 \ldots$ in a partnership, define the set of *partnership histories* at the beginning of the period t, 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 C/D 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, when we write a "strategy distribution" $s \in \mathbf{S}$, we mean the distribution in $\mathcal{P}(\mathbf{S})$ that puts mass one on s.

2.3 Average payoff function

From now on we focus on match-independent strategies. 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**.[5] Stationarity is needed to explicitly compute the average long-run payoff of each strategy.

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, 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[6]). 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

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

To explain, the s-player loses the partner 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.

[5] Because the partnerships may not end simultaneously for all players, the social strategy distribution is not the same as the one in the matching pool, in general. However, there should be a continuous bijection between stationary strategy distributions in the matching pool and those in the society. See footnote 7 in Greve-Okuno.

[6] 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.

Using this, (1) can be arranged as follows.

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

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

$$v(s;p) := (1-\delta)V(s;p) = \frac{\sum_{s' \in supp(p)} p(s') U(s,s')}{\sum_{s' \in supp(p)} p(s') L(s,s')}. \tag{2}$$

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

$$v(s;p) = \frac{U(s;p)}{L(s;p)}. \tag{3}$$

2.4 Fundamentally asymmetric equilibrium

Let us define the *fundamentally asymmetric equilibrium* of Fujiwara-Greve et al. (2015).

Definition 4 Let c_0-*strategy* be a match-independent strategy[7] as follows: in any period $t = 1, 2, \ldots$ of a partnership and after any partnership history, play C and Stay if and only if (C, C) is observed, i.e., for 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 $(a, a') \neq (C, C)$. (The first coordinate in (a, a') is the player's own action.)

Let d_0-*strategy* be a match-independent strategy as follows: in any period $t = 1, 2, \ldots$ and after any partnership history $h \in H_t$, play D and Leave for any observation, i.e., $x_{d_0,t}(h) = D$, $y_{d_0,t}(h, (a, a')) =$ Leave for any $(a, a') \in \{C, D\}^2$.

[7] 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.

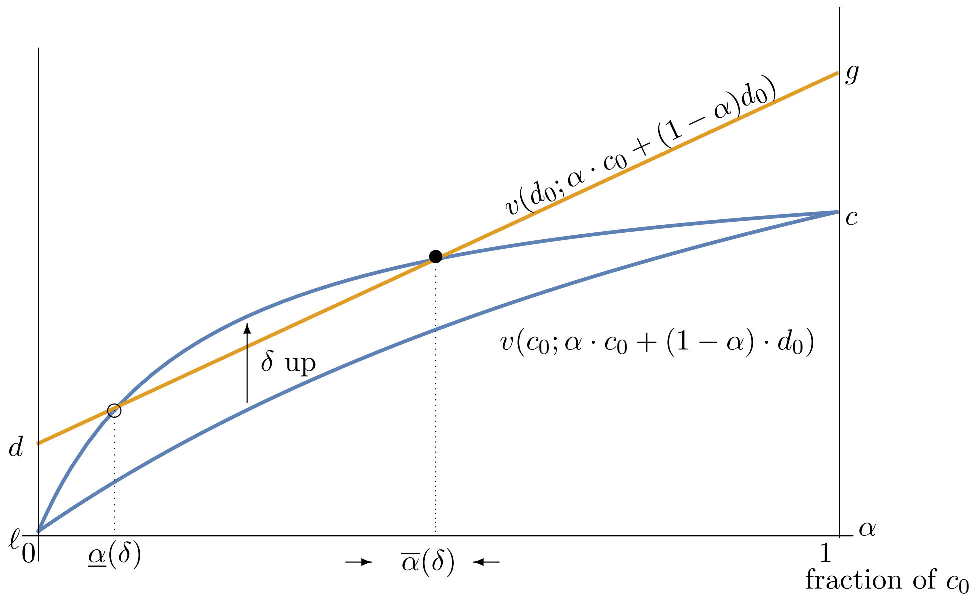

Figure 2: Locally stable c_0-d_0 equilibrium

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[8] for any $\delta \in (0,1)$, but the c_0-strategy does not. An interesting property of the VSRPD is that these two strategies can make a Nash equilibrium.[9]

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 $s' \in \mathcal{S}$,

$$v(s;p) \geqq v(s';p). \tag{4}$$

Lemma 1 *(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 $\overline{\alpha}(\delta) \in (0,1)$, 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: there exists a neighborhood U of $\overline{\alpha}(\delta)$ such that for any $\alpha \in U$,*

$$\alpha \gtrless \overline{\alpha}(\delta) \iff v(d_0; \alpha \cdot c_0 + (1-\alpha) \cdot d_0) \gtrless v(c_0; \alpha \cdot c_0 + (1-\alpha) \cdot d_0). \tag{5}$$

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

[9]Although we did not explicitly derive $v(s';p)$ for each general private strategy 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 illustrates the intuition behind the payoff-equivalence of the c_0- and the d_0-strategy and the local stability, for sufficiently large δ. For notational convenience, for each $z \in \{c_0, d_0\}$ and any $\alpha \in [0,1]$, denote

$$v_0(z;\alpha) := v(z; \alpha \cdot c_0 + (1-\alpha) \cdot d_0).$$

Explicitly, the average payoff of the c_0-strategy is

$$v_0(c_0;\alpha) = \frac{\alpha \cdot U(c_0,c_0) + (1-\alpha)U(c_0,d_0)}{\alpha \cdot L(c_0,c_0) + (1-\alpha)L(c_0,d_0)} = \frac{\alpha \cdot \frac{c}{1-\delta^2} + (1-\alpha)\ell}{\alpha \cdot \frac{1}{1-\delta^2} + (1-\alpha)}. \tag{6}$$

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

$$v_0(d_0;\alpha) = \frac{\alpha \cdot U(d_0,c_0) + (1-\alpha)U(d_0,d_0)}{\alpha \cdot L(d_0,c_0) + (1-\alpha)L(d_0,d_0)} = \alpha \cdot g + (1-\alpha)d. \tag{7}$$

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 $\alpha \in (0,1)$, 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.*

3 Evolutionary Stability of the Fundamentally Asymmetric Equilibrium

3.1 S-Neutral Stability

Fujiwara-Greve et al. (2015) showed that the fundamentally asymmetric equilibrium does not satisfy Greve-Okuno's neutral stability. We rename it as S-Neutral Stability where S signals the focus on symmetric-strategy or single-strategy mutants. (Greve-Okuno focused on monomorphic equilibria with monomorphic mutant distributions, and thus this stability was not so strong.)

Definition 6 *(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 by S-NS($\mathbf{S}$)) if, for any $s' \in \mathbf{S}$, there exists

$\overline{\epsilon} \in (0,1)$ such that for any $s \in supp(p^*)$ and any $\epsilon \in (0, \overline{\epsilon})$,

$$v(s;(1-\epsilon)p^* + \epsilon \cdot s') \geqq v(s';(1-\epsilon)p^* + \epsilon \cdot s'). \tag{8}$$

Remark 1 *(Remark 1 of Fujiwara-Greve et al., 2015) For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, $p_{c_0 d_0}(\delta)$ does not satisfy S-NS($\mathbf{S}$).*

For the motivation of the rest of the paper, we explain the proof. Consider a strategy $D_{c_0} \in \mathbf{S}$ such that, in the first period of any match, the D_{c_0}-strategy plays D and stay for any observation and in the second period of any match, its continuation strategy is c_0 for any observation in $t = 1$. (See Definition 9.) Then for any $\epsilon \in (0,1)$

$$v(D_{c_0};(1-\epsilon)p_{c_0 d_0}(\delta) + \epsilon \cdot D_{c_0}) > v(d_0;(1-\epsilon)p_{c_0 d_0}(\delta) + \epsilon \cdot D_{c_0}).$$

This is because the D_{c_0}-strategy behaves the same way as the d_0-strategy does towards the equilibrium strategies, but cooperate among themselves, i.e., the D_{c_0}-strategy conducts secret-handshake (Robson, 1990 and Matsui, 1991). Therefore $p_{c_0 d_0}(\delta)$ does not satisfy (8).

However, it may be possible that (for sufficiently small ϵ) the c_0-strategy earns a higher fitness than the mutant, and the two equilibrium strategies **together** may beat the mutant as follows[10]:

$$\begin{aligned} &\overline{\alpha}(\delta)v(c_0;(1-\epsilon)p_{c_0 d_0}(\delta) + \epsilon \cdot D_{c_0}) + \{1-\overline{\alpha}(\delta)\}v(d_0;(1-\epsilon)p_{c_0 d_0}(\delta) + \epsilon \cdot D_{c_0}) \\ \geqq\ & v(D_{c_0};(1-\epsilon)p_{c_0 d_0}(\delta) + \epsilon \cdot D_{c_0}). \end{aligned}$$

This mean payoff of the equilibrium pure-strategies is the standard "fitness" measurement of a strategy distribution (e.g., Sandholm, 2010, Izquierdo et al., 2018). However, as Fujiwara-Greve et al. (2015) shows, the above inequality may not hold for all Prisoner's Dilemma payoff parameters.

Another possibility is that the mutants are "dispersed" so that they do not concentrate on the D_{c_0}-strategy. For example, some mutants may choose the D_{d_0}-strategy, which plays D and stay for any observation in the first period, and play the d_0-strategy as the continuation strategy in the second period, for any observation in $t = 1$. (See Definition 9.) This strategy imitates both the d_0- and

[10]For example, consider $(g, c, d, \ell, \delta, \epsilon) = (1250, 1150, 700, 600, 0.9, 0.1)$. Then the C-D ratio is $\overline{\alpha}(\delta) \approx 0.7622$, and the mean payoff of the c_0- and the d_0-strategy in the post-entry distribution is $(0.7622) \cdot 1106 + (0.2377) \cdot 1077.3 \approx 1099.17$, while the payoff of the D_{c_0}-mutant is 1099.03.

the D_{c_0}-strategy and exploits the latter. As we see below, for any Prisoner's Dilemma, there exists $A \subset (0,1)$ such that for any $\alpha \in A$,

$$\begin{aligned}&\overline{\alpha}(\delta) \cdot v(c_0; (1-\epsilon)p_{c_0d_0}(\delta) + \epsilon \cdot q) + \{1 - \overline{\alpha}(\delta)\} v(d_0; (1-\epsilon)p_{c_0d_0}(\delta) + \epsilon \cdot q) \\ \geqq\ & \alpha \cdot v(D_{c_0}; (1-\epsilon)p_{c_0d_0}(\delta) + \epsilon \cdot q) + (1-\alpha) v(D_{d_0}; (1-\epsilon)p_{c_0d_0}(\delta) + \epsilon \cdot q),\end{aligned}$$

where $q = \alpha D_{c_0} + (1-\alpha) D_{d_0}$. That is, the fundamentally asymmetric equilibrium on average performs (weakly) better than a range of mixed-mutant distributions. We generalize the above idea in the next subsection.

3.2 Tolerant Strategies and Neutral Stability

To avoid measure-theoretic complications which do not give us new game-theoretic insights, from now on we restrict our attention to strategy distributions with a countable support and let $\mathcal{Q} \subset \mathcal{P}(\mathbf{S})$ be the set of strategy distributions with a countable support. Moreover, we parameterize the stability definition with respect to the set M of potential mutant distributions.

Definition 7 Given $M \subset \mathcal{Q}$, a stationary strategy distribution in the matching pool $p^* \in \mathcal{Q}$ is *Neutrally Stable with respect to* the mutant distributions from M (denoted by NS(M)) if, for any $q \in M$, there exists $\overline{\epsilon} \in (0,1)$ such that for any $\epsilon \in (0, \overline{\epsilon})$,

$$\sum_{s \in supp(p*)} p^*(s) v(s; (1-\epsilon)p^* + \epsilon \cdot q) \geqq \sum_{s' \in supp(q)} q(s') v(s'; (1-\epsilon)p^* + \epsilon \cdot q). \tag{9}$$

Clearly if p^* satisfies S-NS($\mathbf{S}$), then it satisfies NS($\mathbf{S}$) but may not satisfy NS(M) for some $M \subset \mathcal{Q}(\subset \mathcal{P}(\mathbf{S}))$. Definition 7 is the "matching-pool distribution" version of the *Izquierdo & van Veelen Neutrally Stable Distribution* in Izquierdo et al. (2018), when M is the set of all strategy distributions with a finite support. The larger the M is, the stronger the stability is. However, Remark 1 shows that M cannot be $\mathcal{Q}$ for the fundamentally asymmetric equilibrium.

From now on we restrict possible mutant strategies as *tolerant strategies*, introduced by Fujiwara-Greve and Okuno-Fujiwara (2017) and elaborated by Fujiwara-Greve and Okuno-Fujiwara (2019).

Definition 8 For any $k = 0, 1, 2, \ldots$, any k-period sequence[11] $\mathbf{X} = (X_1, X_2, \ldots, X_k) \in \{C, D\}^k$, and any $z \in \{c_0, d_0\}$, the *k-period tolerant strategy*, denoted $\mathbf{X}_z$, is a (match-independent) strategy such that, in periods $t = 1, 2, \ldots$ of a partnership,

(Tolerant phase)

if $t = 1$: play $X_1 \in \{C, D\}$ and stay for any observation;

if $2 \leqq t \leqq k$: if (C, C) is observed in the previous period, play the c_0-strategy as the continuation strategy, and otherwise play $X_t \in \{C, D\}$ and stay for any observation;

(Commitment phase)

if $t = k + 1$: if (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.

In words, a tolerant strategy has three characteristics: (a) during the tolerant phase, unless (C, C) is established, it stays and plays its "planned action sequence" $\mathbf{X}$ regardless of the partner's reaction, (b) if (C, C) is established during the tolerant phase, it shifts to the c_0-strategy immediately (those that play only D during the tolerant phase can be vacuously interpreted this way as well), and (c) when the tolerant phase is over, it commits to one of the c_0- and the d_0-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. Therefore the fundamentally asymmetric equilibrium does not satisfy S-NS, even if we restrict the set of mutant strategies to the tolerant strategies, since the D_{c_0}-strategy discussed in subsection 3.1 is a tolerant strategy.

One-period tolerant strategies, which help the reader to understand the later results, are as follows.

Definition 9 The C_{c_0}-strategy is a (match-independent) strategy such that

$t = 1$: play C and stay for any observation;

$t = 2$: play the c_0-strategy as the continuation strategy for any observation in $t = 1$.

The C_{d_0}-strategy is a strategy such that

[11] As a convention, let $\{C, D\}^0 = \emptyset$. Hence the c_0- and the d_0-strategies are included as degenerate tolerant strategies.

$t = 1$: play C and stay for any observation;

$t = 2$: play the c_0-strategy if (C, C) is observed in $t = 1$, and d_0-strategy otherwise.

The D_{c_0}-strategy is a strategy such that

$t = 1$: play D and stay for any observation;

$t = 2$: play the c_0-strategy as the continuation strategy for any observation in $t = 1$.

The D_{d_0}-strategy is a strategy such that

$t = 1$: play D and stay for any observation;

$t = 2$: play the d_0-strategy as the continuation strategy for any observation in $t = 1$.

In words, X_{c_0} (resp. X_{d_0}) strategy first plays $X \in \{C, D\}$ in a new match to "take a look" at the new partner before committing to the c_0 (resp. d_0) strategy in the second period, and moreover, if (C, C) is established in the first period, any of these strategies is happy to shift to the c_0-strategy. (D_z-strategies can be trivially interpreted this way.)

Let us justify our focus on the tolerant strategies. First, as argued in Fujiwara-Greve and Okuno-Fujiwara (2019) (and illustrated in Figure 3), the set of tolerant strategies **generates all relevant play paths** in the society, i.e., the only play paths which are not feasible by tolerant strategies are the ones that Leave after (C, C) and those that stay but play D after (C, C). These two classes of play paths generate lower continuation payoffs than that of some tolerant strategy. Alternatively, tolerant-strategy mutants can be interpreted as **potentially equilibrium entrants**: tolerant strategies can constitute a Nash equilibrium if they enter with the C-D ratio (see Fujiwara-Greve and Okuno-Fujiwara, 2019), but the strategies that generate the infeasible paths cannot.[12]

Second, among the tolerant strategies, the **explicit payoff comparison is possible**. In general, it is very difficult to explicitly compare the average payoffs of two (extensive-form) strategies

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

because both the denominator and the numerator can be different, due to the endogenous lengths of partnerships.

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

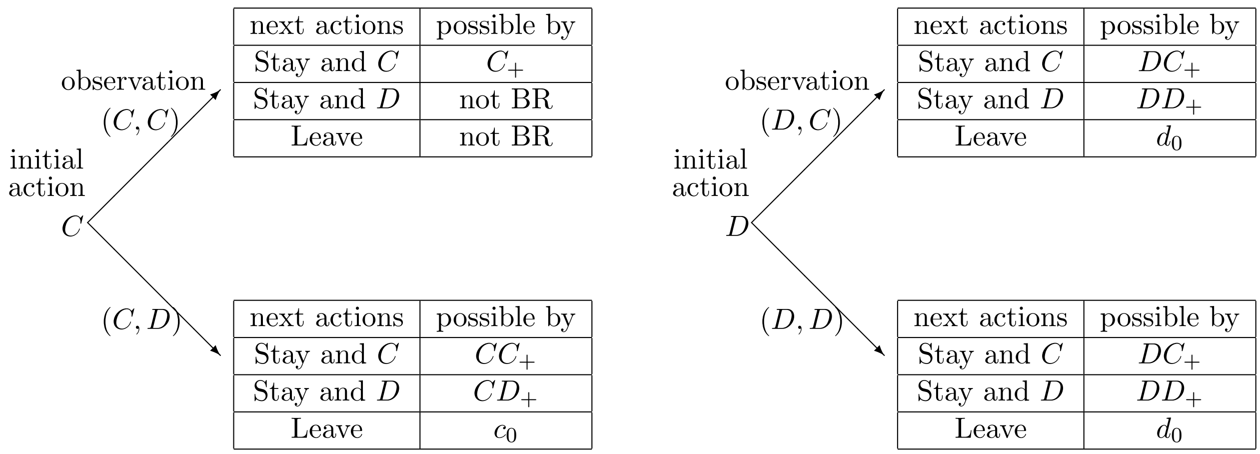

Figure 3: Feasible action sequences over two periods by tolerant strategies

By contrast, the structure of the tolerant strategies is so organized that a k-period tolerant strategy never experiences longer partnerships than a $k+1$-period tolerant strategy with the same initial action plan for $k+1$ periods (see Table 2). This makes the explicit payoff comparison possible, as in Lemmas 5 and 6 (in Appendix).

For each $k = 0, 1, 2, \ldots$, the set of k-period tolerant strategies is denoted by

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

and the set of all tolerant strategies with k-period or longer tolerant phase is denoted by $\tilde{S}_k^{\infty} := \cup_{j=k}^{\infty} \tilde{S}_j$.

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

$$C_+ := \{\mathbf{X}_z \in \tilde{S}_0^{\infty} \mid X_1 = C\}$$

$$D_+ := \{\mathbf{X}_z \in \tilde{S}_0^{\infty} \mid X_1 = D\}.$$

In words, a strategy in C_+ (resp. D_+) is a tolerant strategy which plays C (resp. D) in $t = 1$ in any match, including the c_0-strategy (resp. the d_0-strategy).

Tolerant strategies have a convenient structure such that adding initial action plan to a tolerant strategy makes a new tolerant strategy.

Definition 10 For each $k = 0, 1, 2, \ldots$, any (initial) k-period action sequence $\mathbf{X} \in \{C, D\}^k$, and any strategy $s \in C_+$ (or $s \in D_+$), define the "concatenated" strategy $\mathbf{X}_s$ as a tolerant strategy such that

$t = 1$: play $X_1 \in \{C, D\}$ and stay for any observation;

$2 \leqq t \leqq k$: if (C, C) is observed in the previous period, play the c_0-strategy as the continuation strategy, and otherwise play X_t and stay for any observation;

$t = k+1$: if (C, C) is observed in the previous period, play the c_0-strategy as the continuation strategy, and otherwise play s as the continuation strategy.

Using the above notation, for each $k = 0, 1, 2, \ldots$ and any $\mathbf{X} \in \{C, D\}^k$, define the extended classes of tolerant strategies with the same k-period initial action plan:

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

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

(For $k = 0$, $\emptyset C_+ = C_+$ and $\emptyset D_+ = D_+$.) In words, any strategy in $\mathbf{X}C_+$ (resp. $\mathbf{X}D_+$) plays the initial action sequence $\mathbf{X}$ for k-periods and C (resp. D) in $k+1$-th period of any match unless (C, C) has been established in the first k periods. (The commitment phase is the $k + m + 1$-th period where m is the length of the tolerant phase of s.)

Lemma 2 *(Weak External Stability) For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, define*

$$M_0(\delta) := \Big\{ q \in \mathcal{P}(\tilde{S}^\infty_0) \mid \sup_{\substack{\mathbf{X} \in \{C,D\}^k,\ k=1,2,\ldots \\ q(\mathbf{X}C_+ \cup \mathbf{X}D_+) > 0}} \frac{q\big(\mathbf{X}C_+\big)}{q\big(\mathbf{X}C_+ \cup \mathbf{X}D_+\big)} < \overline{\alpha}(\delta) \Big\}.$$

For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$ and any $q \in M_0(\delta)$, there exists $\overline{\epsilon} \in (0, 1)$ such that, for any $\epsilon \in (0, \overline{\epsilon})$,

$$\forall s' \in C_+ \cap supp(q) \setminus \{c_0\},\ v(c_0; p^{PE}(\epsilon)) \geqq v(s'; p^{PE}(\epsilon));$$

$$\forall s' \in D_+ \cap supp(q) \setminus \{d_0\},\ v(d_0; p^{PE}(\epsilon)) \geqq v(s'; p^{PE}(\epsilon)),$$

where $p^{PE}(\epsilon) = (1 - \epsilon) p_{c_0 d_0}(\delta) + \epsilon \cdot q$.

Proof of Lemma 2. See Appendix.

Lemma 2 shows that any mutant strategy (in $\tilde{S}_1^{\infty}$) has at least one equilibrium strategy which is not worse than the mutant. This property is similar to External Stability in the definition of von Neumann-Morgenstern stable set.

The mutant distributions in $M_0(\delta)$ have the following structure: if $\mathbf{X}_{c_0}$ is present in the mutant distribution q, then $supp(q)$ also contains some strategy from $\{\mathbf{X}_{d_0}, \mathbf{X}D_{c_0}, \mathbf{X}D_{d_0}, \ldots\}$ with the same $\mathbf{X}$ and

$$\frac{q(\mathbf{X}_{c_0}, \mathbf{X}C_{c_0}, \mathbf{X}C_{d_0}, \ldots)}{q(\mathbf{X}_{c_0}, \mathbf{X}_{d_0}, \mathbf{X}C_{c_0}, \mathbf{X}C_{d_0}, \mathbf{X}D_{c_0}, \mathbf{X}D_{d_0} \ldots)} < \overline{\alpha}(\delta).$$

It is also fine that $supp(q) \subset \{\mathbf{X}_{d_0} \mid \exists \mathbf{X} \in \{C, D\}^k, \exists k\}$.

The boundary of $M_0(\delta)$ includes the distribution such that the relative ratio between $\mathbf{X}_{c_0}$-strategy and $\mathbf{X}_{d_0}$-strategy is the C-D ratio, for any $\mathbf{X}$. This is a Nash equilibrium itself (see the companion paper, Fujiwara-Greve and Okuno-Fujiwara, 2019). In view of Remark 1, we cannot hope for weak external stability when mutant distributions are outside of $M_0(\delta)$.

We give an intuition of the proof of Lemma 2 for $s' \in C_+ \setminus \{c_0\}$. Note that s' is an m-period tolerant strategy with the initial action plan $C\mathbf{X} = (C, X_2, X_3, \ldots, X_m)$ (m can be 1, in which case $C\mathbf{X} = C$) and a commitment strategy $z \in \{c_0, d_0\}$. Table 2 shows the payoff sequences of such s' up to two-period tolerant ones, when they meet a partner also using a tolerant strategy.

Remark 2 *Take any $p \in \mathcal{P}(\tilde{S}_0^{\infty})$ and any $z \in \{c_0, d_0\}$.*

$$\begin{aligned} U(C_z;p) - U(c_0;p) &= \delta^2[p(DC_+) \cdot U(z, c_0) + p(DD_+) \cdot U(z, d_0)] \qquad (10)\\ U(CC_z;p) - U(C_{c_0};p) &= \delta^4[p(DDC_+) \cdot U(z, c_0) + p(DDD_+) \cdot U(z, d_0)]\\ U(CD_z;p) - U(C_{d_0};p) &= \delta^4[\sum_{\tilde{X} \in \{C,D\}} p(D\tilde{X}C_+) \cdot U(z, c_0) + p(D\tilde{X}D_+) \cdot U(z, d_0)]. \end{aligned}$$

The proof of Remark 2 is straightforward from Table 2 and is omitted. Intuitively, (a) when we compare two tolerant strategies having the same initial action plan for k-periods, they have the same payoff when they meet a partner using a k'-period tolerant strategy where $k' \leqq k$, and (b) if (C, C) is established with a partner before or in the k-th period, they also have the same payoff. Therefore, to compare the 0-period tolerant c_0-strategy and a 1-period tolerant C_z-strategy (i.e., $k = 1$), they have the same payoff when (a) they meet a 0-period tolerant strategy (c_0 or d_0) or (b) they meet

you\ partner	c_0	C_{c_0}	C_{d_0}	d_0	D_{c_0}	D_{d_0}
CX_z	$c, c, \ldots$	$c, c, \ldots$	$c, c, \ldots$	ℓ	$\ell, \begin{cases} c, c, \ldots & \text{if } X = C \\ g & \text{if } X = D \end{cases}$	$\ell, \begin{cases} \ell & \text{if } X = C \\ d & \text{if } X = D \end{cases}$
C_x	$c, c, \ldots$	$c, c, \ldots$	$c, c, \ldots$	ℓ	$\ell, \begin{cases} c, c, \ldots & \text{if } x = c_0 \\ g & \text{if } x = d_0 \end{cases}$	$\ell, \begin{cases} \ell & \text{if } x = c_0 \\ d & \text{if } x = d_0 \end{cases}$
c_0	$c, c, \ldots$	$c, c, \ldots$	$c, c, \ldots$	ℓ	ℓ	ℓ

	CC_w	CD_w	DC_w	DD_w
CX_z	$c, c, \ldots$	$c, c, \ldots$	$\ell, \begin{cases} c, c, \ldots & \text{if } X = C \\ g, U(z,w) & \text{if } X = D \end{cases}$	$\ell, \begin{cases} \ell, U(z,w) & \text{if } X = C \\ d, U(z,w) & \text{if } X = D \end{cases}$
C_x	$c, c, \ldots$	$c, c, \ldots$	$\ell, \begin{cases} c, c, \ldots & \text{if } x = c_0 \\ g & \text{if } x = d_0 \end{cases}$	$\ell, \begin{cases} \ell & \text{if } x = c_0 \\ d & \text{if } x = d_0 \end{cases}$
c_0	$c, c, \ldots$	$c, c, \ldots$	ℓ	ℓ

Table 2: Payoff sequence comparison of C-start strategies

a C-start tolerant strategy. Their payoff difference thus arises only with the partners in the class $D_+ \setminus \{d_0\} = DC_+ \cup DD+$.

Another key observation from Table 2 is that the payoff of C_z is the same with any partner in $DC_+ = \{D_{c_0}, DC_{c_0}, DC_{d_0}, \ldots\}$ (strategies that play (D, C) for the first two periods). Similarly, C_z gets the same payoff with any partner in DD_+ (strategies that play (D, D) for the first two periods). This is because the C_z-strategy commits to $z \in \{c_0, d_0\}$ in the second period of any match, so that the partner's plan from the third period on is irrelevant. In general, it is the (weakly) earlier-committing strategy that determines the payoff sequence.

Using Remark 2, let us show that, if $q \in M_0(\delta)$, for sufficiently small ϵ's, the post-entry average payoff of the c_0-strategy is weakly greater than that of the C_z-strategy and CX_z-strategy. Take any $q \in M_0(\delta)$ and let $p^{PE}(\epsilon) = (1-\epsilon)p_{c_0 d_0}(\delta) + \epsilon \cdot q$. From (10), for any $z \in \{c_0, d_0\}$ and any $\epsilon \in (0, 1)$,

$$\begin{aligned} \frac{U(C_z; p^{PE}(\epsilon)) - U(c_0; p^{PE}(\epsilon))}{L(C_z; p^{PE}(\epsilon)) - L(c_0; p^{PE}(\epsilon))} &= \frac{\delta^2[p^{PE}(DC_+) \cdot U(z, c_0) + p^{PE}(DD_+) \cdot U(z, d_0)]}{\delta^2[p^{PE}(DC_+) \cdot L(z, c_0) + p^{PE}(DD_+) \cdot L(z, d_0)]} \\ &= \frac{q(DC_+) \cdot U(z, c_0) + q(DD_+) \cdot U(z, d_0)}{q(DC_+) \cdot L(z, c_0) + q(DD_+) \cdot L(z, d_0)}. \end{aligned}$$

If $q(DC_+, DD_+) > 0$, then the above equality and $q \in M_0(\delta)$ imply that

$$\frac{U(C_z; p^{PE}(\epsilon)) - U(c_0; p^{PE}(\epsilon))}{L(C_z; p^{PE}(\epsilon)) - L(c_0; p^{PE}(\epsilon))} = v_0\Big(z; \frac{q(DC_+)}{q(DC_+, DD_+)}\Big) < v_0(z; \overline{\alpha}) = v_0(c_0; \overline{\alpha}).$$

Note that $v_0(z; \frac{q(DC_+)}{q(DC_+, DD_+)})$ is independent from ϵ. Since

$$\lim_{\epsilon \to 0} v(c_0; p^{PE}(\epsilon)) = v_0(c_0; \overline{\alpha}),$$

there exists $\overline{\epsilon} \in (0, 1)$ such that for any $\epsilon \in (0, \overline{\epsilon})$,

$$v_0\Big(z; \frac{q(DC_+)}{q(DC_+, DD_+)}\Big) < v(c_0; p^{PE}(\epsilon)).$$

This translates into the following:

$$\begin{aligned}
&\frac{U(C_z; p^{PE}(\epsilon)) - U(c_0; p^{PE}(\epsilon))}{L(C_z; p^{PE}(\epsilon)) - L(c_0; p^{PE}(\epsilon))} = v_0\Big(z; \frac{q(DC_+)}{q(DC_+, DD_+)}\Big) < v(c_0; p^{PE}(\epsilon)) = \frac{U(c_0; p^{PE}(\epsilon))}{L(c_0; p^{PE}(\epsilon))} \\
\iff & L(c_0; p^{PE}(\epsilon))[U(C_z; p^{PE}(\epsilon)) - U(c_0; p^{PE}(\epsilon))] < U(c_0; p^{PE}(\epsilon))[L(C_z; p^{PE}(\epsilon)) - L(c_0; p^{PE}(\epsilon))] \\
\iff & \frac{U(C_z; p^{PE}(\epsilon))}{L(C_z; p^{PE}(\epsilon))} < \frac{U(c_0; p^{PE}(\epsilon))}{L(c_0; p^{PE}(\epsilon))} \\
\iff & v(C_z; p^{PE}(\epsilon)) < v(c_0; p^{PE}(\epsilon)).
\end{aligned}$$

If $q(DC_+, DD_+) = 0$, then clearly for any $\epsilon \in (0, 1)$,

$$v(C_z; p^{PE}(\epsilon)) = v(c_0; p^{PE}(\epsilon)).$$

The comparison of the C_{c_0}-strategy and a CC_z-strategy is similar and, for sufficiently small ϵ's, we have

$$v(CC_z; p^{PE}(\epsilon)) \leqq v(C_{c_0}; p^{PE}(\epsilon)).$$

Finally, to compare C_{d_0} and CD_z, we use the fact that

$$\frac{A_1 + A_2}{B_1 + B_2} \leqq \max\{\frac{A_1}{B_1}, \frac{A_2}{B_2}\},$$

for any $B_1, B_2 > 0$ (see Lemma 4). From Remark 2 and by the above inequality, if at least one of $q(DCC_+, DCD_+) > 0$ or $q(DDC_+, DDD_+) > 0$ holds,

$$\frac{U(CD_z; p^{PE}(\epsilon)) - U(C_{d_0}; p^{PE}(\epsilon))}{L(CD_z; p^{PE}(\epsilon)) - L(C_{d_0}; p^{PE}(\epsilon))} \leqq \max\Big\{v_0\Big(z; \frac{q(DCC_+)}{q(DCC_+, DCD_+)}\Big), v_0\Big(z; \frac{q(DDC_+)}{q(DDC_+, DDD_+)}\Big)\Big\}.$$

Then $q \in M_0(\delta)$ implies that

$$\frac{U(CD_z; p^{PE}(\epsilon)) - U(C_{d_0}; p^{PE}(\epsilon))}{L(CD_z; p^{PE}(\epsilon)) - L(C_{d_0}; p^{PE}(\epsilon))} < v_0(c_0; \overline{\alpha}).$$

For sufficiently small ϵ's, $v(c_0; p^{PE}(\epsilon))$ is close to $v_0(c_0; \overline{\alpha})$ so that

$$\frac{U(CD_z; p^{PE}(\epsilon)) - U(C_{d_0}; p^{PE}(\epsilon))}{L(CD_z; p^{PE}(\epsilon)) - L(C_{d_0}; p^{PE}(\epsilon))} < v(c_0; p^{PE}(\epsilon)) = \frac{U(c_0; p^{PE}(\epsilon))}{L(c_0; p^{PE}(\epsilon))}$$

holds so that

$$v(CD_z; p^{PE}(\epsilon)) < v(C_{d_0}; p^{PE}(\epsilon)).$$

If both $q(DCC_+, DCD_+) = 0$ and $q(DDC_+, DDD_+) = 0$, then clearly

$$v(CD_z; p^{PE}(\epsilon)) = v(C_{d_0}; p^{PE}(\epsilon)).$$

Hence, combining the weak inequalities, for sufficiently small ϵ's,

$$\forall X \in \{C, D\},\ v(CX_z; p^{PE}(\epsilon)) \leqq v(c_0; p^{PE}(\epsilon)).$$

This logic generalizes for an arbitrary C-start (resp. D-start) mutant strategy in comparison with the c_0-strategy (resp. d_0-strategy).

Moreover, the local stability of the fundamentally asymmetric equilibrium implies Neutral Stability.

Proposition 1 *For any $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$, the fundamentally asymmetric equilibrium $p_{c_0 d_0}(\delta) = \overline{\alpha}(\delta) c_0 + \{1 - \overline{\alpha}(\delta)\} d_0$ is NS($M_0(\delta)$).*

Proof of Proposition 1. Fix an arbitrary $\delta \in (\underline{\delta}_{c_0 d_0}, 1)$ and take any $q \in M_0(\delta)$. For each $\epsilon \in (0, 1)$, let $p^{PE}(\epsilon) := (1 - \epsilon)[\overline{\alpha}(\delta) \cdot c_0 + \{1 - \overline{\alpha}(\delta)\} d_0] + \epsilon \cdot q$. Then for each $z \in \{c_0, d_0\}$, $v(z; p^{PE}(\epsilon)) = v_0(z; (1-\epsilon)\overline{\alpha}(\delta) + \epsilon \cdot q(C_+))$, that is, the average payoff of z-strategy depends only on the share of the set C_+.

Lemma 2 implies that there exists $\tilde{\epsilon} \in (0, 1)$ such that for any $\epsilon \in (0, \tilde{\epsilon})$,

$$\forall s' \in C_+ \cap supp(q) \setminus \{c_0\},\ v(c_0; p^{PE}(\epsilon)) \geqq v(s'; p^{PE}(\epsilon)); \tag{11}$$

$$\forall s' \in D_+ \cap supp(q) \setminus \{d_0\},\ v(d_0; p^{PE}(\epsilon)) \geqq v(s'; p^{PE}(\epsilon)). \tag{12}$$

By (11) and (12), the group-mean payoff of q is bounded as follows.

$$\begin{aligned} \forall \epsilon \in (0, \tilde{\epsilon}),\ & \sum_{s' \in supp(q)} q(s') v\big(s'; p^{PE}(\epsilon)\big) \\ & \leqq q(C_+) v\big(c_0; p^{PE}(\epsilon)\big) + \{1 - q(C_+)\} v\big(d_0; p^{PE}(\epsilon)\big). \end{aligned} \tag{13}$$

In view of (13), it suffices to show that there exists $\hat{\epsilon} \in (0,1)$ such that for any $\epsilon \in (0, \hat{\epsilon})$,

$$
\begin{aligned}
& q(C_+)v\big(c_0; p^{PE}(\epsilon)\big) + \{1 - q(C_+)\}v\big(d_0; p^{PE}(\epsilon)\big) \\
\leqq\ & \overline{\alpha}(\delta)v\big(c_0; p^{PE}(\epsilon)\big) + \{1 - \overline{\alpha}(\delta)\}v\big(d_0; p^{PE}(\epsilon)\big) \\
=\ & \sum_{z \in supp(p_{c_0 d_0})} p_{c_0 d_0}(s) v(s; p^{PE}(\epsilon)).
\end{aligned} \tag{14}
$$

For any $\epsilon \in (0,1)$, define

$$
\begin{aligned}
f(\epsilon) &:= \big[\overline{\alpha}(\delta)v(c_0; p^{PE}(\epsilon)) + \{1 - \overline{\alpha}(\delta)\}v(d_0; p^{PE}(\epsilon))\big] - \big[q(C_+)v(c_0; p^{PE}(\epsilon)) + \{1 - q(C_+)\}v(d_0; p^{PE}(\epsilon))\big] \\
&= [\overline{\alpha}(\delta) - q(C_+)]\Big\{v(c_0; p^{PE}(\epsilon)) - v(d_0; p^{PE}(\epsilon))\Big\}.
\end{aligned}
$$

Notice that for any $\epsilon \in (0,1)$,

$$
\overline{\alpha}(\delta) \gtreqless p^{PE}(\epsilon)(C_+) = (1-\epsilon)\overline{\alpha}(\delta) + \epsilon q(C_+) \iff \overline{\alpha}(\delta) \gtreqless q(C_+).
$$

By the local stability (5), there exists $\hat{\epsilon} > 0$ such that for any $\epsilon \in (0, \hat{\epsilon})$, $p^{PE}(\epsilon)(C_+) \in U$ and

$$
\overline{\alpha}(\delta) - q(C_+) \gtreqless 0 \iff \overline{\alpha}(\delta) \gtreqless p^{PE}(\epsilon)(C_+) \iff v(c_0; p^{PE}(\epsilon)) - v(d_0; p^{PE}(\epsilon)) \gtreqless 0.
$$

(See Figure 2.) Therefore, $f(\epsilon) \geqq 0$, i.e., (14), holds for any $\epsilon \in (0, \hat{\epsilon})$. Finally, let $\overline{\epsilon} = \min\{\tilde{\epsilon}, \hat{\epsilon}\}$. Then

$$
\forall \epsilon \in (0, \overline{\epsilon}), \sum_{s' \in supp(q)} q(s')v\big(s'; p^{PE}(\epsilon)\big) \leqq \sum_{z \in supp(p_{c_0 d_0})} p_{c_0 d_0}(s) v(s; p^{PE}(\epsilon)).
$$

□

4 Concluding Remarks

We have provided a new evolutionary foundation to the fundamentally asymmetric equilibrium, consisting of conditional cooperators and myopic defectors. Although we restricted the potential set of mutants, we allow any Prisoner's Dilemma.

The stability of the fundamentally asymmetric equilibrium is not so weak because the monomorphic d_0-distribution, which is a Nash equilibrium for any $\delta \in (0,1)$, does not satsify NS($M_0(\delta)$) for sufficiently large δ. To see this, note that there is $q \in M_0(\delta)$ such that $q = \alpha \cdot D_{c_0} + (1-\alpha)D_{d_0}$ with the property $\alpha < \overline{\alpha}$. In this distribution, the mutant D_{d_0}-strategy uses the other mutant strategy

D_{c_0} as the "spring board" to earn more than d on average, while the d_0-strategy cannot. Moreover, when δ is large, D_{c_0}-strategy also earns more than d on average, because the partnership with another D_{c_0}-strategy lasts for a very long time.

Remark 3 *There exists $\underline{\delta} \in [\underline{\delta}_{c_0 d_0}, 1)$ such that for any $\delta \in (\underline{\delta}, 1)$, the monomorphic distribution d_0 does not satisfy NS($M_0(\delta)$).*

Proof. Consider a mutant distribution of the form $q = \alpha \cdot D_{c_0} + (1-\alpha) D_{d_0}$. For each $\epsilon \in (0,1)$, write $p^{PE}(\epsilon) = (1-\epsilon) \cdot d_0 + \epsilon \cdot q$. The average post-entry payoffs are as follows.

$$\begin{aligned}
&v(d_0; p^{PE}(\epsilon)) = d;\\
&v(D_{d_0}; p^{PE}(\epsilon)) = \frac{(1-\epsilon)d + \epsilon[\alpha(d+\delta^2 g) + (1-\alpha)(d+\delta^2 d)]}{(1-\epsilon) + \epsilon(1+\delta^2)} > d;\\
&v(D_{c_0}; p^{PE}(\epsilon)) = \frac{(1-\epsilon)d + \epsilon[\alpha(d+\delta^2 \frac{c}{1-\delta^2}) + (1-\alpha)(d+\delta^2 \ell)]}{(1-\epsilon) + \epsilon[\frac{\alpha}{1-\delta^2} + (1-\alpha)(1+\delta^2)]}.
\end{aligned} \tag{15}$$

Let us compare (15) with $v(d_0; p^{PE}(\epsilon))$.

$$\begin{aligned}
&v(D_{c_0}; p^{PE}(\epsilon)) > v(d_0; p^{PE}(\epsilon))\\
\iff & d + \epsilon\delta^2\alpha\frac{c}{1-\delta^2} + \epsilon\delta^2(1-\alpha)\ell > d[1 + \epsilon\delta^2\alpha\frac{1}{1-\delta^2} + \epsilon\delta^2(1-\alpha)]\\
\iff & \alpha(c-d) > (1-\alpha)(1-\delta^2)(d-\ell).
\end{aligned}$$

Let $x := \delta^2$. The above inequality is equivalent to

$$\alpha > \frac{(1-x)(d-\ell)}{c-d+(1-x)(d-\ell)} =: h(x).$$

By differentiation, the threshold $h(x)$ is decreasing in x (and converges to 0 as $x \to 1$):

$$h'(x) = -\frac{(c-d)(d-\ell)}{[c-d+(1-x)(d-\ell)]^2} < 0.$$

Note also that $\overline{\alpha}(\delta)$ is increasing in δ.[13] Thus there exists $\delta^* \in (0,1)$ such that for any $\delta \in (\delta^*, 1)$,

$$h(x) < \overline{\alpha}(\delta).$$

[13] This can be seen from the fact that $v_0(d_0;\alpha)$ is constant in δ and $\frac{\partial v_0(c_0;\alpha)}{\partial \delta} = \frac{2(1-\alpha)\alpha\delta(c-\ell)}{\{1-\delta^2(1-\alpha)\}^2} > 0$, which is proved in the Proof of Proposition 1 of Fujiwara-Greve et al. (2015).

Let $\underline{\delta} = \max\{\delta^*, \underline{\delta}_{c_0 d_0}\}$. Then for any $\delta \in (\underline{\delta}, 1)$, there exists $\alpha < \overline{\alpha}(\delta)$ such that

$$v(D_{c_0}; p^{PE}(\epsilon)) > v(d_0; p^{PE}(\epsilon)),$$

so that $q \in M_0(\delta)$ and for any $\epsilon \in (0, 1)$,

$$\sum_{s' \in supp(q)} q(s')v(s'; (1-\epsilon) \cdot d_0 + \epsilon \cdot q) > v(d_0; (1-\epsilon) \cdot d_0 + \epsilon \cdot q).$$

$\square$

Finally, we note that when $g - c$ is very small and δ is large, $\overline{\alpha}(\delta)$ becomes close to 1. To see this, let $\delta \to 1$. Then $v_0(c_0; \alpha) \to c$. Since $v_0(d_0; \alpha) = \alpha g + (1-\alpha)d$ is independent of δ, we have

$$\lim_{\delta \to 1} \overline{\alpha}(\delta) = \frac{c-d}{g-d}.$$

Hence for "small stake" Prisoner's Dilemmas with very long-lived players, the sufficient mutant set $M_0(\delta)$ is not restrictive, and the fundamentally asymmetric equilibrium consists mostly of the c_0-players. That is, nearly-symmetric cooperative society is neutrally stable with respect to nearly-all tolerant strategy distributions.

Appendix: Proof of Lemma 2

Before proving Lemma 2, we give two technical lemmas to make it easy to compare the average payoffs of different strategies, when 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 the matching pool and two pure strategies $s, \hat{s} \in supp(p)$, 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 in the society and are easy to compute.

Lemma 3 *For any p and any $s, \hat{s} \in supp(p)$,*

(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 of Lemma 3. By definition,

$$\begin{aligned} & v(s;p) > v(\hat{s};p) \\ \iff & \frac{U(s;p)}{L(s;p)} > \frac{U(s;p) + \Delta U(\hat{s}, s; p)}{L(s;p) + \Delta L(\hat{s}, s; p)} \\ \iff & U(s;p) \cdot \Delta L(\hat{s}, s; p) > L(s;p) \cdot \Delta U(\hat{s}, s; p), \end{aligned} \tag{16}$$

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 (16) 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 (16) is equivalent to

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

$\square$

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 4 *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 of Lemma 4. 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,

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

□

Next, we give bounds to the average payoff difference between a k-period tolerant strategy in C_+ (resp. D_+) and the c_0-strategy (resp. the d_0-strategy), by generalizing Remark 2.

We introduce the notion of *induced* strategies.

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

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

Lemma 5 *For any $k = 1, 2, \ldots$, take any k-period tolerant strategy $\mathbf{X}_z \in \tilde{S}_k$ and its induced $(k-1)$-period tolerant strategy $\mathbf{X}'_{x_k} \in \tilde{S}_{k-1}$. Then, for any $p \in \mathcal{P}(\tilde{S}_0^\infty)$, their 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 5. 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:

$$\forall \tau \leqq k-1, \ \forall \ \tilde{\mathbf{X}}_{\tilde{z}} \in \tilde{S}_\tau, \ U(\hat{s}, \tilde{\mathbf{X}}_{\tilde{z}}) - U(s^*, \tilde{\mathbf{X}}_{\tilde{z}}) = 0. \tag{17}$$

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. $\square$

Lemma 6 *For 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 $p \in \mathcal{P}(\tilde{S}_0^\infty)$,*

$$\frac{\Delta U(\mathbf{X}_z, \mathbf{X}'_{x_{m+1}}; p)}{\Delta L(\mathbf{X}_z, \mathbf{X}'_{x_{m+1}}; p)} \leqq \max_{j=m+1,\ldots,k} \max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C), \forall t=1,\ldots,j}} v\Big(x_{j+1}; \frac{p(\tilde{\mathbf{X}}C_+)}{p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big); \tag{18}$$

$$\frac{\Delta U(\mathbf{X}_z, \mathbf{X}'_{x_{m+1}}; p)}{\Delta L(\mathbf{X}_z, \mathbf{X}'_{x_{m+1}}; p)} \geqq \min_{j=m+1,\ldots,k} \min_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C), \forall t=1,\ldots,j}} v\Big(x_{j+1}; \frac{p(\tilde{\mathbf{X}}C_+)}{p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big), \tag{19}$$

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 6. 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 5, for each $j = m+1, m+2, \ldots, k$,

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

(Note that $x_{k+1} = z$.) By Lemma 4, 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.

$$\frac{\Delta U(s_j, s_{j-1}; p)}{\Delta L(s_j, s_{j-1}; p)} \leqq \max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,j}} v\Big(x_{j+1}; \frac{p(\tilde{\mathbf{X}}C_+)}{p(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big). \tag{20}$$

Furthermore, by Lemma 4 again,

$$\frac{\Delta U(s_k, s_m; p)}{\Delta L(s_k, s_m; p)} = \frac{\sum_{j=m+1}^{k} \Delta U(s_j, s_{j-1}; p)}{\sum_{j=m+1}^{k} \Delta U(s_j, s_{j-1}; p)} \leqq \max_{j=m+1,\ldots,k} \frac{\Delta U(s_j, s_{j-1}; p)}{\Delta L(s_j, s_{j-1}; p)}. \tag{21}$$

(20) and (21) imply that

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

This is (18). Derivation of (19) is analogous. $\square$

Proof of Lemma 2. Fix an arbitrary $q \in M_0(\delta)$ and let $p^{PE}(\epsilon) = (1-\epsilon)p_{c_0 d_0} + \epsilon q$. Recall that, for each $s \in \{c_0, d_0\}$,

$$\lim_{\epsilon \to 0} v(s; p^{PE}(\epsilon)) = v(c_0; \overline{\alpha}(\delta)).$$

Hence, in view of Lemma 3 (i), it suffices to prove that, for any $\epsilon \in (0,1)$,

$$\forall s' \in C_+ \cap supp(q) \setminus \{c_0\},\ \frac{\Delta U(s', c_0; p^{PE}(\epsilon))}{\Delta L(s', c_0; p^{PE}(\epsilon))} < v(c_0; \overline{\alpha}(\delta)); \tag{22}$$

$$\forall s' \in D_+ \cap supp(q) \setminus \{d_0\},\ \frac{\Delta U(s', d_0; p^{PE}(\epsilon))}{\Delta L(s', d_0; p^{PE}(\epsilon))} < v(d_0; \overline{\alpha}(\delta)). \tag{23}$$

Without loss of generality, assume that $C_+ \cap supp(q) \setminus \{c_0\} \neq \emptyset$ and take any $s' \in C_+ \cap supp(q) \setminus \{c_0\}$. There exists k such that $s' \in \tilde{S}_k$. The bound (18) in Lemma 6 implies that

$$\frac{\Delta U(s', c_0; p^{PE}(\epsilon))}{\Delta L(s', c_0; p^{PE}(\epsilon))} \leqq \max_{j=1,\ldots,k} \max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,j}} v\Big(x_{j+1}; \frac{p^{PE}(\tilde{\mathbf{X}}C_+)}{p^{PE}(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big),$$

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

$$\frac{\Delta U(s', c_0; p^{PE}(\epsilon))}{\Delta L(s', c_0; p^{PE}(\epsilon))} \leqq \max_{j=1,\ldots,k} \max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,j}} v\Big(x_{j+1}; \frac{q(\tilde{\mathbf{X}}C_+)}{q(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big),$$

Thus the upper bound is independent of ϵ.

Since $v(c_0; \alpha)$ and $v(d_0; \alpha)$ are increasing in α, and by the assumption $q \in M_0(\delta)$, for each $j = 1, \ldots, k$,

$$\max_{\substack{\tilde{\mathbf{X}} \in \{C,D\}^j \\ (X_t, \tilde{X}_t) \neq (C,C)\ \forall t=1,\ldots,j}} v\Big(x_{j+1}; \frac{q(\tilde{\mathbf{X}}C_+)}{q(\tilde{\mathbf{X}}C_+ \cup \tilde{\mathbf{X}}D_+)}\Big) < v(x_{j+1}; \overline{\alpha}(\delta)) = v(c_0; \overline{\alpha}(\delta)).$$

Therefore (22) holds. The case of $s' \in D_+ \cap supp(q) \setminus \{d_0\}$ is analogous. $\square$

References

Carmichael, L., MacLeod, B. 1997. Gift Giving and the Evolution of Cooperation. Int. Econ. Rev. 38(3), 485-509.

Datta, S. 1996. Building trust. Mimeo. London School of Economics.

Duffie, D., Sun, Y. 2012. The exact law of large numbers for independent random matching. J. Econ. Theory 147 (3), 1105-1139.

Duffie, D., Quao, L., Sun, Y. 2018. Dynamic directed random matching. J. Econ. Theory 174, 124-183.

Eeckhout, J. 2006. Minorities and endogenous segregation. Rev. Econ. Stud. 73(1), 31-53.

Ellison, G. 1994. Cooperation in the Prisoner's Dilemma with Anonymous Random Matching. Rev. Econ. Stud. 61(3), 567–588.

Fujiwara-Greve, T. and Okuno-Fujiwara, M. 2009. Voluntarily separable repeated prisoner's dilemma. Rev. Econ. Stud. 76(3), 993–1021.

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

Fujiwara-Greve, T., Okuno-Fujiwara, M. 2019. Tolerance and behavioral diversity. Mimeo. Keio University.

Fujiwara-Greve, T., Okuno-Fujiwara, M. and Suzuki, N. 2015. Efficiency may improve when defectors exist. Econ. Theory 60(3), 423–460.

Garcìa, J. and van Veelen, M. 2016. In and out of equilibrium I: Evolution of strategies in repeated games with discounting. J. Econ. Theory 161, 161-189.

Ghosh, P. and Ray, D. 1996. Cooperation in community interaction without information flows. Rev. Econ. Stud. 63(3), 491-519.

Heller, Y. 2017. Instability of belief-free equilibria. J. Econ. Theory 168, 261-286.

Immorlica, N., Lucier, B., and Rogers, B. 2014. Cooperation in anonymous dynamic social networks. Mimeo. Microsoft Research and Washington University in St. Louis.

Izquierdo, Luis, Izquierdo, Segismundo, and van Veelen, Matthijs. 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. J. Econ. Dyn. Control 46, 91–113.

Kandori, M. 1992. Social Norms and Community Enforcement. Rev. Econ. Stud. 59(1), 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. J. Law, Econ. & Org. 12(1), 214-233.

McAdams, D. 2011. Performance and turnover in a stochastic partnership. Am. Econ. J.:Micro. 3(4), 107-142.

Matsui, A. 1991. Cheap talk and cooperation in a society. J. Econ. Theory 54(2), 245-258.

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

Maynard Smith, J. and Price, G. 1973. The logic of animal conflict. Nature 246,15-18.

Rob, R. and Yang, H. 2010. Long-term relationships as safeguards. Econ. Theory 43(2), 143-166.

Robson, A. 1990. Efficiency in evolutionary games: Darwin, Nash and the secret handshake. J. Th. Bio. 144(3), 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.

Selten, R. 1983. Evolutionary stability in extensive two-person games. Math. Soc. Sci. 5(3), 269-363.

Sun, Y. 2006. The exact law of large numbers via Fubini extension and characterization of insurable risks. J. Econ. Theory 126(1), 31-69.

Swinkels, J. 1992. Evolutionary Stability with Equilibrium Entrants. J. Econ. Theory 57, 306–332.

van Veelen, M. 2012. Robustness against indirect invasions. Games Econ. Behav. 74(1), 382-393.