Stability of an Efficient Equilibrium in Voluntarily Separable Repeated Prisoner's Dilemma*

NOBUE SUZUKI

Abstract

Stability of the most efficient Nash equilibrium in the VS-RPD is re-examined, with the notion that mutation of a strategy would invoke that of another. It is shown that, under simultaneous mutation of such strategies, the basin of attraction of efficient Nash equilibrium will be considerably large. However, the efficient Nash equilibrium is not locally stable unless there is some fixed cost for playing complicated mutant strategies.

Keywords: voluntarily separable, Prisoner's Dilemma, cooperation, VSRPD
JEL classification: C 73

*The author is grateful to Takako Fujiwara-Greve and Masahiro Okuno-Fujiwara for useful comments. This work was supported by JSPS Grant-in-Aid for Scientific Research (C), No.21530171. The usual disclaimer applies.

1 Introduction

Modern economic relationships, such as Internet transactions, often take the form that strangers meet randomly to trade or collaborate and continue the relationship by mutual agreement. Ordinary repeated-game framework does not apply to such situations because it assumes that players repeat the stage game for certain periods without an option to terminate.

Recently, a line of research of voluntarily separable repeated games (VSRG) has appeared.[1] In a VSRG, players are randomly matched and repeat the stage game only by mutual agreement; each player has a unilateral power to end the relationship and find a new partner in a random matching pool. When the stage game is a two-action Prisoner's Dilemma, Fujiwara-Greve and Okuno-Fujiwara (2009), hereafter abbriviated as GO (2009), showed that there are evolutionary stable distributions consisting of 'trust-building strategies': for certain periods at the beginning of a new match, players do not cooperate but continue the partnership ('trust-building phase'), and after that the players cooperate and keep the partnership if and only if the partner also cooperated ('cooperation phase').

Okuno-Fujiwara et al. (2007) and GO (2009) have shown that, under sufficiently large discount factor, trust-building strategy c_1 can constitute a neutrally stable monomorphic equilibrium. Suzuki (2008) further investigated stability of such trust-building monomorphic equilibrium. A trust-building strategy not only is locally stable, but also can invade to the bimorphic Nash equilibrium consisting of the C-trigger-with-quit strategy c_0 and hit-and-run strategy d_0.

However, a situation in which all mutants playing c_1 would invoke mitation of another strategy, say d_1. d_1 player keeps the partnership

[1]See, for example, Datta (1996), Ghosh and Ray (1996), Carmichael and MacLeod (1997), Kranton (1996), and Fujiwara-Greve and Okuno-Fujiwara, (2009).

at the end of the first period to cheat c_1 player in the second period. In this paper, I will investigate the effect of simultaneous mutation of c_1 and d_1 to the c_0-d_0 equilibrium.

It is shown that c_1 earns less payoff when c_1 and d_1 simultaneously mutates, compared with the case in which every mutant is c_1: the basin of attraction of the c_0-d_0 equilibrium drastically expands. However, this is not sufficient to make c_0-d_0 equilibrium locally stable. c_1 and d_1 strategies are more complex than c_0 and d_0 strategies. If playing c_1 or d_1 is assosiated with some memory cost, then the c_0-d_0 equilibrium is locally stable.

2 The Model

Consider a large society of a continuum of players with measure 1. The time is discrete. At the end of each period, each player exits from the society for an exogenous reason (which we call a "death") with probability $1 - \delta$ $(0 < \delta < 1)$. If a player dies, a new player enters the society, keeping the population size constant. In each period, players without a partner (including the newly born players) enter the random matching pool and form pairs to play the *Voluntarily Separable Repeated Prisoner's Dilemma* as follows.

Randomly matched players first play the Prisoner's Dilemma (see Table 1) by choosing action C or D simultaneously. The actions in the Prisoner's Dilemma are observable only by the partners. After that, based on the observation, each player chooses whether to keep the partnership (action k) or end it (action e) simultaneously. If at least one player chooses e, the partnership ends and both players (if they survive) go to the random matching pool in the next period. If both chose action k, unless one of them dies, they play the Prisoner's Dilemma together again in the next period, skipping the matching process. If the partner dies, the surviving player goes to the random

P1 \ P2	C	D
C	$c,\ c$	$\ell,\ g$
D	$g,\ \ell$	$d,\ d$

Table 1: Prisoner's Dilemma

matching pool in the next period.

The payoff in a period is determined only by the action profile in the Prisoner's Dilemma, as in Table 1. Assume that $g > c > d > \ell$ and $2c \geq g+\ell$. The latter is for simplicity and to make the symmetric action profile (C, C) efficient.

The game continues with probability δ from an individual player's point of view. Thus we focus on the expected total payoff, with δ being the effective discount factor of a player.

As in GO (2009) and Suzuki (2008), I shall assume throughout the current paper that the actions within a partnership are observable perfectly between the partners but not observable by any other player. Therefore newly matched players have no information about the past actions of each other.

Let $t = 1, 2, \ldots$ denote the periods in the current partnership, rather than the calendar time. For each t, define $H_t := [\{C, D\}^2]^{(t-1)}$ as the set of partnership histories at the beginning of the t-th period of a partnership. Since the result of the Prisoner's Dilemma in the t-th period is observable before players proceed to the second stage, the set of partnership history at the beginning of the second stage in the t-th period is $H_t \times \{C, D\}^2 := [\{C, D\}^2]^t$. If it is a new match $(t = 1)$, then the partnership history is $H_1 = \{\emptyset\}$ at the beginning of the period and $H_1 \times \{C, D\}^2 := \{C, D\}^2$ at the beginning of the second stage of that period.

3 Strategies and Expected Payoff

Let $V(s, s')$ be the expected total payoff during a match given to strategy s matched with s', discounted according to the probability of two players surviving. The expected length of this match will be denoted by $L(s, s')$. Let $p(s)$ be the share of strategy s under population distribution p. Then, with strategy distribution $p \in \mathcal{P}$, the expected per period payoff of $s \in S$ is given by

$$v(s; p) = \frac{\sum_{s' \in S} p(s')V(s, s')}{\sum_{s' \in S} p(s')L(s, s')}. \tag{1}$$

Under the assumptions on information sepecified in the previous section, it is natural to consider the strategies in which action is based only upon the partnership history. The following two categories of rules of thumb, c_T and d_T, are examples of such strategies.

- c_T-strategies:
  - (i) for $t \leq T$, play (D, k) no matter what is observed; and
  - (ii) for $t > T$, play C in the Prisoner's Dilemma and choose k if and only if its observed outcome in that period is (C, C).
- d_T-strategies:
  - (i) for $t \leq T$, play (D, k); and
  - (ii) for $t > T$, play (D, e).

c_0-strategy start cooperating from the first period of a match and end the partnership as soon as the opponent plays d. This can be attributed to be an out-for-tat strategy, or to be analogous to the C-tit-for-tat-with-quit strategy in the repeated game framework. On the other hand, d_0 is a typical defective strategy which plays (D, e) in the first period.

c_T-strategies with $T \geq 1$ are called trust-building strategies in GO (2009) and in other papers on the VSRPD. These strategies start with defection and simply keeps the partnership for the first T period(s) of match. For example, c_1-player plays D in the first period of a match and switch to c_0-strategy from the second period. Non-cooperation in the first period allow c_1-players to avoid being sacrificed by d_0-players. This is why c_1-strategy exhibited substantial robustness in the three-strategy dynamics of c_0-d_0-c_1 in GO (2009) and Suzuki (2008).

More precisely, consider a situation in which c_1-strategy is mutating to a Nash equilibrium p_{cd} consisting only of c_0 and d_0 strategies. When c_1-strategy meets the incumbent strategies, it behaves as if it is d_0 and hence best responds to p_{cd}. On the other hand, c_1-strategy better responds to c_1-strategy than the incumbent strategies do, provided that the discount factor δ is large enough and that the population distribution is sufficiently close to p_{cd}.

Of course, if mutation of c_1 is associated with an increase in $p(d_0)$ relative to $p(c_0)$, d_0 and c_1 becomes relatively less fit and thus c_1 might not be able to invade. Suzuki (2008) investigated the full dynamics of the three-strategy distribution of c_0-d_0-c_1. It has been shown that, as depicted in Figure 1, the basin of attraction of p_{cd} is fairly small and, starting from most initial distribution of c_0-d_0-c_1, the steady state will be characterized by a predominance of c_1.

However, this argument still ignores the possibility that mutation of c_1-strategy not only changes the incumbent distribution, but also invoke mutation of other strategies. From a defective player's viewpoint, if she is sure that the opponent is playing c_1, she had better to play d_1 rather than d_0: she can earn defective gain in the second period by keeping the partnership at the end of the first period. Although robustness of c_1-strategy to such cheating strategies is analyzed in Suzuki (2008), the effect of simultaneous mutation of c_1 and d_1 strategies to c_0-d_0 distribution remained unexamined.

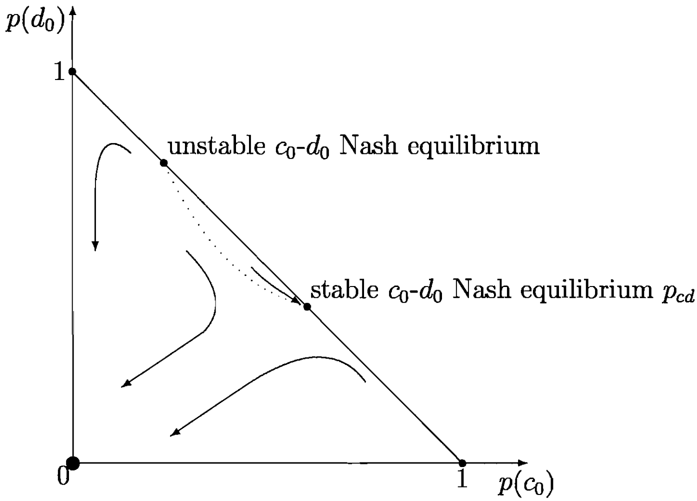


Figure 1: Phase diagram for c_0-d_0-c_1

4 Four-Strategy Distribution

Consider four-strategy distribution of c_0, d_0, c_1 and d_1 in the VSRPD. Dealing with the global dynamics of these four srtategies will make the analysis unnecessarily difficult. Recall that the current purpose is to examine the robustness of c_0-d_0 Nash equilibrium against mutation of c_1 and d_1 strategies. Therefore, the local analysis around p_{cd} is the key issue here.

Suzuki (2008) shows that p_{cd} is locally stable with respect to the two-strategy dynamics of c_0-d_0: d_0-players earn higher payoff when there are too many c_0-players to give g to them; and less when there are too many d_0-players to give d. Moreover, as it will be explained in more detail later, c_0-strategy is indifferent between meeting d_0, c_1 or d_1: in any case, c_0-player gets ℓ in the first period and terminate the partnership.

Given these, local robustness of p_{cd} to mutation of c_1 and d_1 strategies can be analyzed in the follwing two steps, as described in Figure 2:

(i) the population share of c_0 strategy will be fixed at the share of c_0 in p_{cd}, namely,

$$p(c_0) = p_{cd}(c_0) = p^*; \tag{2}$$

and

(ii) with $p(c_0)$ kept at p^*, analyze the dynamics between the three strategies d_0, c_1 and d_1.

If the dynamics of (ii) exhibits convergence to d_0, then it implies convergence to p_{cd} from the initial distribution of c_0-d_0-c_1-d_1. On the other hand, if c_1 or d_1 evolves in the dynamics of (ii), then both c_0 and d_0 are less fit than c_1 or d_1: such distribution is not in the basin of attraction of p_{cd}.

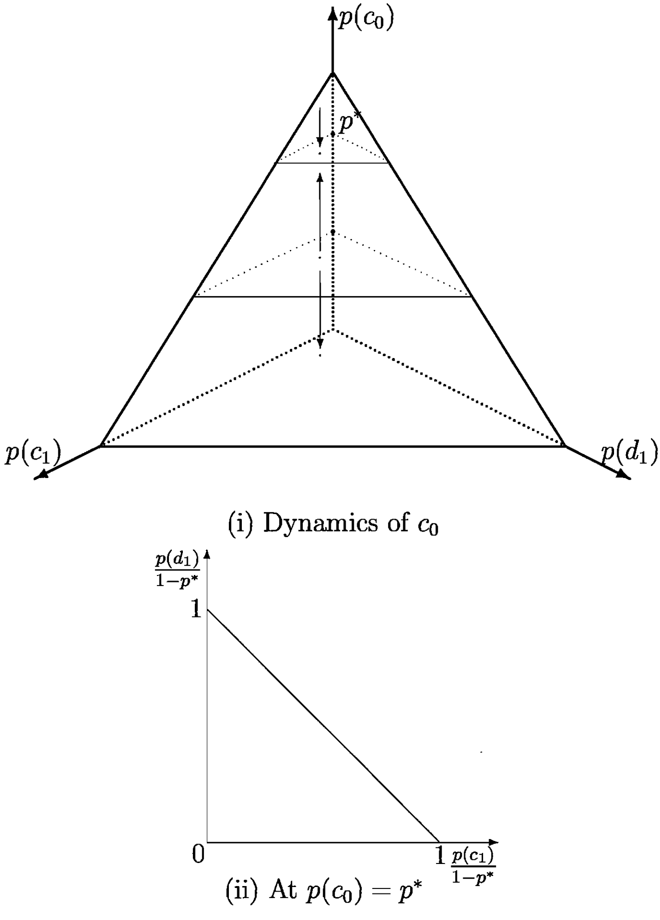


Figure 2: Dealing with c_0-d_0-c_1-d_1 Dynamics

For notational simplisity, the population share of strategies other than p^* will be written as

$$\begin{aligned} p(c_1) &= q & (3)\\ p(d_0) &= 1-p^*-q-u=r & (4)\\ p(d_1) &= u & (5)\end{aligned}$$

where $0 \leq q, r, u \leq 1$.

Table 2 demonstrates the payoff flow within the partnership of row player matched with a column player. Using this table, the average (per period) payoff of each strategy $v(\cdot;p)$ under strategy distribution $p=(p^*,q,r,u)$ is calculated as:

$$\begin{aligned} v(c_0;p) &= \frac{p^*\frac{c}{1-\delta^2}+(1-p^*)\ell}{\frac{p^*}{1-\delta^2}+1-p^*}\\ &= d+\frac{p^*\frac{c-d}{1-\delta^2}+(1-p^*)(\ell-d)}{1+\frac{p^*\delta^2}{1-\delta^2}} & (6)\\ v(c_1;p) &= \frac{p^*g+q\left[d+\frac{\delta^2 c}{1-\delta^2}\right]+rd+u(d+\delta^2\ell)}{p^*+\frac{q}{1-\delta^2}+r+u(1+\delta^2)}\\ &= d+\frac{p^*(g-d)+\frac{q\delta^2(c-d)}{1-\delta^2}+u\delta^2(\ell-d)}{1+\frac{q\delta^2}{1-\delta^2}+u\delta^2} & (7)\\ v(d_0;p) &= p^*g+(1-p^*)d=d+p^*(g-d) & (8)\\ v(d_1;p) &= \frac{p^*g+q(d+\delta^2 g)+rd+u(1+\delta^2)d}{p^*+q(1+\delta^2)+r+u(1+\delta^2)}\\ &= d+\frac{p^*(g-d)+q\delta^2(g-d)}{1+(q+u)\delta^2} & (9)\end{aligned}$$

Assume that dynamics of (ii) is governed by the following two

opponent (probability)	c_0 (p^*)	c_1 (q)	d_0 (r)	d_1 (u)
c_0	$c, c, \ldots$	ℓ	ℓ	ℓ
c_1	g	$d, c, c, \ldots$	d	d, ℓ
d_0	g	d	d	d
d_1	g	d, g	d	d, d

Table 2: Payoff in Each Match

equations:[2]

$$\dot{q} = (v(c_1; p) - v(c_0; p))q \tag{10}$$

and

$$\dot{u} = (v(d_1; p) - v(c_0; p))u. \tag{11}$$

It turns that

$$\dot{q} \geq 0 \quad \text{iff} \quad p^* \leq \frac{q}{u} \tag{12}$$

and

$$\dot{u} \geq 0 \quad \text{iff} \quad p^* \leq \frac{q}{u}. \tag{13}$$

That is, if the share of c_1 among the mutant is less than the share of c_0 in the original c_0-d_0 Nash equilibrium, then the mutant strategies are weaker than the incumbent strategies. More intuilively, successfull invasion of c_1 requires that a substantial fraction of those who keep the partnership after (D, D) in the first period start cooperating from the second period. Otherwise, c_1 players gets ℓ more often in the second period. d_1 players are also worse off when there are too many d_1, because they get d more often.

Since p^* tends to be considerably large, this implies that the basin of attraction of p_{cd} in the dynamics of (ii) is also large. However,

[2]There is no substantial change in the result when the usual replicator dynamics is used.

it cannot be said that p_{cd} is locally stable: it is not robust against mutation of a tiny fraction of c_1-d_1 with $\frac{q}{u} \geq p^*$.

Figure 3 demonstrates the dynamics (ii) for two different sets of parametor values. In either case, p_{cd} has a large basin of attraction but it is not locally stable.

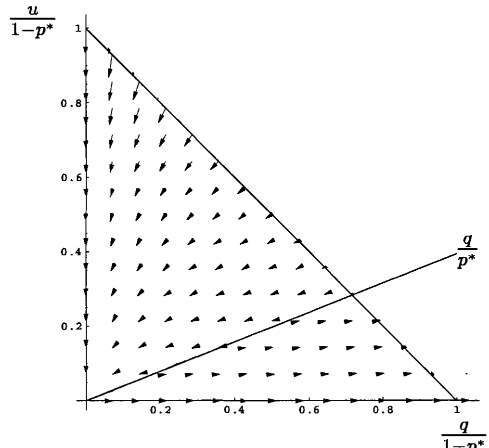
5 Concluding Remarks

Okuno *et al.* (2009) shows that the Nash equilibrium consisting only of c_0 and d_0 strategies is the most efficient. But, nevertheless, GO (2009) and Suzuki (2008) claimed that such c_0-d_0 equilibrium to be unstable, since it will be invaded by c_1-strategy.

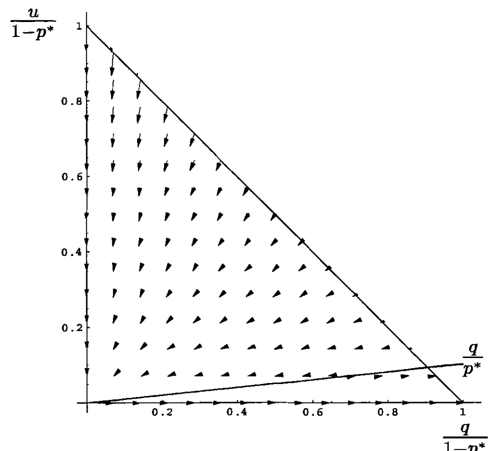
The current paper takes into account of the possibility that mutation of c_1 invokes mutation of d_1. When c_1 and d_1 simultaneously mutate, c_0-d_0 equilibrium has considerably large basin of attraction, because c_1 will be wiped out if there are sufficiently large fraction of d_1. However, this is not sufficient to make the c_0-d_0 equilibrium locally stable.

It would be argued that c_1 and d_1 strategies are more complicated than c_0 and d_0 strategies are. Therefore, it seems natural to assume that players need to pay some memory cost to play such complicated strategies correctly. If a fixed cost m is incurred to those who are playing c_1 or d_1, then the c_0-d_0 equilibrium is locally stable. Figure 4 demonstrates this for two different sets of parametor values.

Another possibility of making the c_0-d_0 equilibrium locally stable is that, as in Okuno *et al.* (2009), mutation is player's strategic experimentation based upon socially shared belief of off-path outcome. No individual tries c_1 if everyone believes the opponent will not cooperate after observing history of defection in the first period.

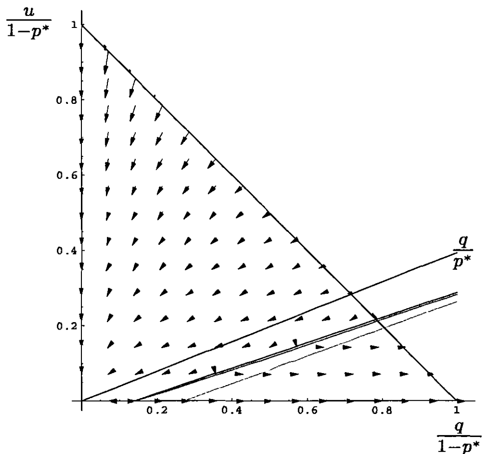


(a) $g = 1.2; c = 1; d = 0; \ell = -1; \delta^2 = .81$

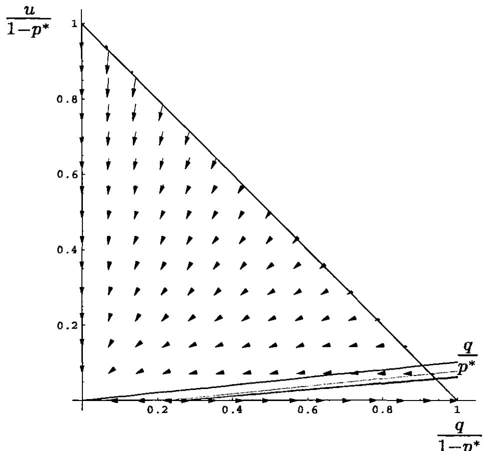


(b) $g = 60; c = 55; d = 0; \ell = -1; \delta^2 = (.95)^2$

Figure 3: Dynamics (ii)



(a) $g = 1.2; c = 1; d = 0; \ell = -1; \delta^2 = .81; m = .02$



(b) $g = 60; c = 55; d = 0; \ell = -1; \delta^2 = (.95)^2; m = .1$

Figure 4: Dynamics (ii) with Memory Cost

References

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

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

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

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

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

Okuno-Fujiwara, M., T. Fujiwara-Greve, and N. Suzuki (2007) "Social Norms and Voluntary Cooperation," *Keizai Kenkyu* (in Japanese), 58(2): 110-121.

Okuno-Fujiwara, M., T. Fujiwara-Greve, and N. Suzuki (2009) "Voluntarily Separable Repeated Prisoner's Dilemma with Shared Belief," presented at the 2nd Taiwan-Dutch and International Game Theory Conference in Taipei; under revision.

Suzuki, N. (2008): "Stability of Cooperative Strategies in the Voluntarily Separable Repeated Prisoner's Dilemma," *The Economic Review of Komazawa University* (in Japanese), 39(3), 55-85.