Third person enforcement in a prisoner's dilemma game

Tatsuhiro Shichijo

June 14, 2012

1 Introduction

Three players play a specific repeated game, in which the stage game is the prisoner's dilemma game illustrated in Table 1, where $P = 45, S = 10, T = 100, R = 75$. In the first stage, players X_1 and X_2 play the prisoner's dilemma game as the stage game. From the second stage, players M and X_1 play the stage game with probability 1/2. Players M and X_2 also play the stage game with probability 1/2. That is, player M plays the stage game with certainty. Players X_1 and X_2 play the stage game with probability 1/2. They play the game an infinite number of times, with a discount factor of $\delta = 3/4$. We assume that each player observes only the outcome of the stage game that s/he plays. For example, player M cannot see the action profile of the first stage.

Players X_1 and X_2 play the stage game against each other only once. However, it is possible that they both play C in the first stage because a third person, player M, may enforce cooperation. We analyze whether both players X_1 and X_2 play C in a sequential equilibrium.

Kandori (1992) showed that a contagious strategy constitutes a cooperative equilibrium in a private monitoring setting if the discount factor is sufficiently large. In section 3 of this paper, we show that the Kandori (1992)-type contagious strategy cannot constitute a cooperative equilibrium under the parameter settings of this paper. In section 4, we show that another type of strategy profile constitutes a sequential equilibrium.

	C	D
C	R	S
D	T	P

Table 1: Prisoner's Dilemma Game

2 Notation

We denote by (a_1a_2) the outcome of the first stage in which player X_1 plays a_1 and player X_2 plays a_2, where $a_1, a_2 \in \{C, D\}$. From the second stage, either player X_1 or player X_2 is chosen to play the stage game. To identify the selected player, we denote by $(X_i a_i a_M)$ the stage t outcome in which player X_i is selected and plays a_i and player M plays a_M. For example, (X_1CD) is the stage outcome in which player X_1 plays C and player M plays D. We denote by $(a_1^1 a_2^1; X_i a_i^2 a_M^2; \dots; X_j a_j^t a_M^t)$ the history of the outcome up to stage t. Let H^t be the set of histories up to stage t. The behavioral strategy of player X_i at stage t depends on the history up to stage $t-1$. The behavioral strategy at stage t of player X_i is described by the function $\sigma_i^t : H^{t-1} \to \{C, D\}$. By contrast, the behavioral strategy of player M depends on who the opponent is. The behavioral strategy of player M is described by the function $\sigma_M^t : H^{t-1} \times \{X_1, X_2\} \to \{C, D\}$. When no player is specified, we use z. For example, $(X_zDD; X_zDD)$ means that one of the following outcomes occurs: $(X_1DD; X_1DD)$, $(X_1DD; X_2DD)$, $(X_2DD; X_1DD)$, $(X_2DD; X_2DD)$. When no action is specified, we use Z. For example, (X_1ZZ) means that one of the following outcomes occurs: $(X_1CC), (X_1CD), (X_1DC), (X_1DD)$. We denote the sequence of σ_i^t by σ_i; i.e., $\sigma_i = (\sigma_i^1, \sigma_i^2, \dots)$. We denote the sequence of σ_M^t by σ_M. Let $\sigma = (\sigma_1, \sigma_2, \sigma_M)$.

3 The contagious strategy

In this section, we show that a Kandori (1992)-type contagious strategy cannot constitute a sequential equilibrium under the parameter settings of this paper. A player who plays a contagious strategy plays D if her/his opponent has previously played D against her/him. For example, if player X_1 plays D against player M, then player M plays D not only against player X_1, but also against player X_2. If s/he has previously played D against a player, then s/he again plays D against that player. [1] For example, if player M plays D against player X_1, then player M uses strategy D against player X_1. Otherwise, s/he plays C.

When player M observes a deviation by her/his opponent, s/he assumes that the deviation occurred in the first stage if it is a reasonable deviation. Suppose player M observed $(X_1CC; X_2DC)$ in the second and third stages. There are two explanations: (i) player X_1 or player X_2 played D in the first stage, but player X_1 played C even though X_1 was supposed to play D; (ii) there was no deviation in the first or second stage, but player X_2 deviates in the third stage for the first time. We assume that player M follows (i) and that the player M uses D against both other players.

[1] Because our game setting is different from that of Kandori (1992), the behavioral strategy is slightly different from Kandori (1992)'s contagious strategy. That is, even if s/he has played D against player X_1, s/he plays C against player X_2.

We can constitute a sequence of assessments with completely mixed strategies that is consistent with the contagious strategy profile and with beliefs that satisfy the above principle. Let γ be the mixed strategy such that strategy C is played with probability 1/2 and strategy D is played with probability 1/2. Let $\hat{\sigma}$ be the contagious strategy. We can base a complete mixed strategy $\tilde{\sigma}$ on $\hat{\sigma}$ with $\epsilon > 0$ as follows:

$$\begin{aligned}
\tilde{\sigma}_i^1(\emptyset) &= (1-\epsilon)\hat{\sigma}_i^1(\emptyset) + \epsilon\gamma \text{ for } i \in \{1,2\} \\
\tilde{\sigma}_i^t(CD; \dots) &= (1-\epsilon)\hat{\sigma}_i^t(CD; \dots) + \epsilon\gamma \text{ for } i \in \{1,2\}, t > 1 \\
\tilde{\sigma}_i^t(DC; \dots) &= (1-\epsilon)\hat{\sigma}_i^t(DC; \dots) + \epsilon\gamma \text{ for } i \in \{1,2\}, t > 1 \\
\tilde{\sigma}_M^t(ZZ; \dots \mid X_i) &= (1-\epsilon)\hat{\sigma}_M^t(ZZ; \dots \mid X_i) + \epsilon\gamma \text{ for } i \in \{1,2\}, t > 1 \\
\tilde{\sigma}_i^t(CC; \dots) &= (1-\epsilon^{1/\epsilon})\hat{\sigma}_i(CC; \dots) + \epsilon^{1/\epsilon}\gamma \text{ for } i \in \{1,2\}, t > 1.
\end{aligned}$$

We can base the belief μ_ϵ on $\tilde{\sigma}$ by Bayes' rule. Taking the limit as $\epsilon \to 0$, $\tilde{\sigma}$ converges to $\hat{\sigma}$ and μ_ϵ converges to a belief that satisfies the above principle. This is because $\lim_{\epsilon\to 0} \epsilon^{1/\epsilon}/\epsilon^k = 0$ for all $k \in \mathbb{N}$.

The payoff of player X_i from the contagious strategy profile is $R+\delta R/(2(1-\delta))$. If player X_i plays D in every stage, her/his payoff is $T+\delta T/2+\delta^2 P/(2(1-\delta))$. If $\delta \geq 0.752903$, $R+\delta R/(2(1-\delta)) \geq T+\delta T/2+\delta^2 P/(2(1-\delta))$. If $\delta = 3/4 = 0.75$, which is the parameter setting in this paper, the contagious strategy profile cannot be a sequential equilibrium.

4 A cooperative equilibrium

In this section, we consider a variation of the contagious strategy and show that the new strategy profile σ constitutes a sequential equilibrium. As with the contagious strategy, this strategy is to play D forever if s/he observed that one of his/her opponents deviated from the strategy. For example, player M plays D against X_1 if player M observed that player X_2 deviated from the strategy profile. The difference between our strategy and the contagious strategy relates to the behavioral strategy in the third stage. If player X_i is selected in the second stage and player $X_j (j \neq i)$ is selected in the third stage, then player X_j is allowed to play strategy D in the third stage . In this case, the outcome in the third stage is $(X_j DC)$. Thereafter, player M and player X_j continue to choose C. On the other hand, if player X_i is selected in the second and third stages, then player X_i must play C in the third stage. In this case, if player X_i plays D in the third stage, then player M plays D thereafter.

For example, $(CC; X_1CC; X_1CC; X_zCC; X_zCC; X_zCC; \dots)$ or $(CC; X_1CC; X_2DC; X_zCC; X_zCC; \dots)$ are outcomes on the path of the strategy.

Formal definitions of σ are as follows:

Definition 1.

$$\sigma_1^1(\emptyset) = \sigma_2^1(\emptyset) = C$$
$$\sigma_1^2(CC) = \sigma_2^2(CC) = C$$
$$\sigma_M^2(ZZ \mid X_i) = C \textit{ for } i \in \{1,2\}$$
$$\sigma_i^3(CC; X_iCC) = C \textit{ for } i = \{1,2\}$$
$$\sigma_i^3(CC; X_jCC) = D \textit{ for } i,j = \{1,2\}, \textit{ where } j \neq i$$
$$\sigma_M^3(ZZ; X_zCC \mid X_z) = C$$
$$\sigma_M^3(ZZ; X_iCD \mid X_j) = C \textit{ for } i,j = \{1,2\}, \textit{ where } j \neq i$$
$$\sigma_i^4(CC; X_iCC; X_iCC) = C \textit{ for } i,j = \{1,2\}, \textit{ where } i \neq j$$
$$\sigma_i^4(CC; X_iCC; X_jZZ) = C \textit{ for } i,j = \{1,2\}, \textit{ where } i \neq j$$
$$\sigma_i^4(CC; X_jZZ; X_iDC) = C \textit{ for } i,j = \{1,2\}, \textit{ where } i \neq j$$
$$\sigma_i^4(CC; X_jZZ; X_jZZ) = C \textit{ for } i,j = \{1,2\}, \textit{ where } i \neq j$$
$$\sigma_M^4(ZZ; X_iCC; X_iCC \mid X_i) = C \textit{ for } i = \{1,2\}$$
$$\sigma_M^4(ZZ; X_iCZ; X_iCZ \mid X_j) = C \textit{ for } i,j = \{1,2\}, \textit{ where } i \neq j$$
$$\sigma_M^4(ZZ; X_iCC; X_jZZ \mid X_i) = C \textit{ for } i,j = \{1,2\}, \textit{ where } i \neq j$$
$$\sigma_M^4(ZZ; X_iCZ; X_jDC \mid X_j) = C \textit{ for } i,j = \{1,2\}, \textit{ where } i \neq j \quad .$$

The behavioral strategy played up to stage 4, which is not listed above, is D. From the fifth stage, the strategy is the same as the contagious strategy. That is, if a player plays D after the fifth stage, her/his opponent subsequently plays D and s/he subsequently plays D against the opponent.

When player M observes a deviation by her/his opponent, as in section 3, player M presumes that this deviation occurred in the first stage if it is reasonable. On the other hand, if player M observes $(X_1CC; X_2CC)$, s/he does not assume that the deviation occurred in the first stage because it is unreasonable. As in section 3, a belief that satisfies the above principle is the limit of the beliefs based on the complete mixed strategy.

We show that the above strategy profile constitutes a sequential equilibrium for $\delta = 0.75$.

Theorem 1. *σ constitutes a sequential equilibrium if $\delta = 0.75$, $P = 45, S = 10, T = 100, R = 75$.*

Proof. We investigate the following cases.

Case 1 (in which (X_iDD) is assumed to be played in the strategy profile σ): The stage payoff obtained from playing C is lower than that obtained from playing D. Regardless of the action s/he takes, the opponent continues to play D in subsequent stages. Playing C never improves the payoff obtained from the next stage. Thus, there is no incentive to deviate.

Case 2 (in which there is no deviation and in which $(X_i CC)$ is assumed to be played in the strategy profile in the second stage or later): The expected continuation payoff obtained by player X_i from playing C is $R+\delta R/(2(1-\delta)) = 187.5$. The expected continuation payoff obtained by player X_i from playing D is, at most, $T + \delta P/(2(1-\delta)) = 167.5$. Thus, player X_i has no incentive to deviate. It is easily checked that the same applies for player M.

Case 3 (in which, after playing $(CC; X_i CC)$, player $X_j (j \neq i)$ is selected to play in the third stage): Clearly, player X_j has no incentive to deviate because player X_j is expected to play D. Player M's expected continuation payoff from playing C is $S+\delta R/(1-\delta) = 235$. Player M's expected continuation payoff from playing D is, at most, $P + \delta/2 \times (P+R)/(1-\delta) = 225$. Thus, there is no incentive to deviate.

Case 4 (in which player M gets to choose an alternative in the second stage): The expected continuation payoff for player M from playing C is $R + \delta R/2 + \delta S/2 + \delta^2 R/(1-\delta) = 275.625$. The expected continuation payoff for player M from playing D is, at most, $T + \delta(S/2 + P/2) + \delta P/(2(1-\delta)) + \delta R/(2(1-\delta)) = 255.625$. Thus, there is no incentive to deviate.

Case 5 (in which the first stage outcome is (CD) and the current stage outcome is assumed to be $(X_i DC)$): The expected continuation payoff for player X_i from playing D is $T+\delta P/(2(1-\delta)) = 167.5$. The payoff obtained from playing C is, at most, $R + \delta T/2 + \delta^2 P/(2(1-\delta)) = 163.125$. Note that player M adopts a type of contagious strategy. Player M plays D against X_i after player M plays against $X_j (j \neq i)$. If player X_i continues to play C whenever player M does not play against $X_j (\neq X_i)$, the expected continuation payoff is:

$$R + \delta \frac{R}{2} + \frac{1}{2} \sum_{s=2} \delta^s \left\{ \left(\frac{1}{2}\right)^{s-1} R + (1 - \left(\frac{1}{2}\right)^{s-1}) P \right\}$$

$$= R + \delta \frac{R}{2} + \frac{1}{2} \left(\frac{\delta^2 (R-P)}{2-\delta} + \frac{\delta^2 P}{1-\delta} \right) = 160.5.$$

Thus, there is no incentive to deviate in this case.

Case 6 (in which player X_i gets to choose an alternative in the first stage): If player X_i plays C, the expected continuation payoff is $R + \delta R/2 + \delta^2 (R + T)/4 + \delta^3 R/(2(1-\delta)) = 191.016$. If player X_i plays D in the first stage, the expected continuation payoff is, at most, $T + \delta T/2 + \delta^2 P/(2(1-\delta)) = 188.125$. Thus, there is no incentive to deviate in this case.

The above results show that our strategy profile constitutes a sequential equilibrium strategy.

□

References

Kandori, Michihiro, “Social Norms and Community Enforcement,” *Review of Economic Studies*, January 1992, *59* (1), 63–80.