Sustaining Cooperation in the Repeated Prisoner's Dilemma with Local Interaction

Xue Jun
The Pennsylvania State University

November 27, 2002

Abstract

This paper studies the repeated prisoner's dilemma in a local interaction setup. We construct a sequential equilibrium in pure strategies that sustains cooperation for sufficiently patient players. The notion of sequential equilibrium is extended to extensive form games with infinite time horizon. The strategy is embedded in an explicitly defined expectation system, which may also be viewed as a finite state automaton. The belief system is derived by perturbing the strategy appropriately.

1 Introduction

This paper considers a society where individuals interact with others locally, but a social norm of cooperative behavior is nonetheless sought to be sustained in society as a whole. The development and stability of social norms of cooperation is usually studied through an infinitely repeated prisoner's dilemma, and we adopt the same approach in this paper.

An example might clarify the nature of the problem. Consider a typical road in residential suburban America where each house has a lawn in front. Each houseowner derives a utility v from a well maintained lawn, but can only see her own lawn and those of her neighbors. The cost c of maintaining one's lawn is strictly greater than v. In the case each homeowner has two neighbors, her payoff, if both neighbors maintain their lawns and she doesn't, is $2v$, while if she does is $3v - c$. If neither neighbor maintains his lawn, her payoff is $v - c$ if she does and 0 otherwise.

Would we expect to see the lawns well maintained along the road in the absence of police enforcement of regulations? This paper argues that this is possible in pure strategies. In general, we want to model a situation in which local interaction doesn't create an intrinsic barrier to global coordination.

1.1 The model

The model has the following features:

1. A straight line (finite or infinite in one or both directions).

2. On each integer point lives an agent. Each agent has two neighbors except the end player(s). Let $N(j)$ denote j's neighborhood.

3. Time is discrete, $t = 1, 2, ..., \infty$.

4. The stage game is the prisoner's dilemma.

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

where $g > 0$, $l > 0$.

5. In each period, each player plays the stage game with his two neighbors. He plays the same action against the two neighbors, and his stage game payoff is the sum of his payoffs against both neighbors.

6. Everybody has the same discount factor δ.

7. Each agent only observes what happens in his own neighborhood.

8. Let h_j^t denote player j's private history: $h_j^t = \left(a_{j-1}^s,\ a_j^s,\ a_{j+1}^s\right)_{s=1}^{t-1}$, where $a_k^s \in \{C,\ D\}, k \in N(j)$[1]. Let H_j denote the set of j's private histories. A pure strategy of player j is a mapping from H_j to $\{C,\ D\}$.

We are looking for a sequential equilibrium[2] that supports cooperation on the line. The first natural candidate is, of course, the trigger strategy[3]. A trigger strategy plays cooperation after any history in which no one in the neighborhood has ever defected, and defection otherwise. It turns out that when $l \geq 1$, the trigger strategy works if $\delta \geq \frac{g}{1+g}$. When $l < 1$, the trigger strategy also works if $\delta \in \left[\frac{g}{1+g},\ \frac{g+l}{1+g}\right]$.

The trigger strategy fails when $l < 1$ and δ is large enough for a simple reason. When the cost of being defected upon is smaller than the payoff to

[1] In this paper sometimes k represents a generic player in $N(j)$, like it does here; sometimes it represents a fixed neighbor of j.

[2] The standard notion of sequential equilibrium is defined for finite extensive form games. We have an extension of it in the next section.

[3] The discussion on trigger strategy and mixed strategy below is taken from a note of Bhaskar.

mutual cooperation, and when people are sufficiently patient, they have an incentive to block the spread of defection. That is, people would rather live between a good neighbor and a bad one, than to punish the bad neighbor and then live between two bad neighbors forever. The problem with the trigger strategy is that punishment is both too severe and too lenient. It is too severe in that even the slightest mistake will never be forgiven; it is too lenient in that further deviation (like blocking) is not further punished.

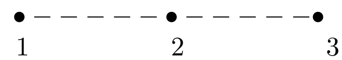
There is a quick fix to the trigger strategy when $l < 1$ and δ is too large. We know that the trigger strategy still works for $\delta \in \left[\frac{g}{1+g}, \frac{g+l}{1+g}\right]$. By Lemma 2 in Ellison(1994), for any sufficiently large δ, we can always use trigger strategy to support cooperation by diluting the original game into a certain number of replica games. Players play the trigger strategy in each replica, and when they play in some replica, they ignore observations from other replicas. Effectively dilution reduces players' discount factor so that it falls back into the interval $\left[\frac{g}{1+g}, \frac{g+l}{1+g}\right]$. The problem with dilution is that it is not uniform with respect to δ: the larger δ is, the more replicas are needed to make people less patient.

An alternative to resolving the problem is using a mixed strategy. The idea is to create some uncertainty about whether one's neighbor is going to punish or block defection, in such a way that this player himself is indifferent between punishing and blocking. The nice feature of the mixed strategy equilibrium is that bad behavior can be localized so that most part of the society is left unaffected even if some player always plays defection. A mixed strategy equilibrium of this form exists when the line is infinite in both directions. With a finite number of agents, there is an "endpoint" problem, to be explained in the next subsection.

The main result of this paper is the following. In the repeated prisoner's dilemma on the finite or infinite line, for any $l < 1$, $g > 0$, there is a sequential equilibrium in pure strategies that supports cooperation for any sufficiently large δ. As in standard folk theorem type of strategies, we need a book keeping device to keep track of punishments and rewards as the game goes on. In standard theory this is done by a machine. Here I do it by defining a pair of expectation operators, $\left(E_j\left(\cdot|h_j^t\right), \;\; E_j\left(E_k\left(\cdot\right)|h_j^t\right)\right)$. For any private history h_j^t, $E_j\left(\cdot|h_j^t\right)$ is the expectation that j forms on the future path of play in his neighborhood, $E_j\left(E_k\left(\cdot\right)|h_j^t\right)$ is the expectation that j forms on his neighbor k's expectation on j's future actions. After h_j^t, j simply does what $E_j\left(\cdot|h_j^t\right)$ expects him to do.

1.2 An example

Let's illustrate the strategy by the following three player expample.



Initially everybody expects everybody to play cooperation forever. Then if say player 2 is surprised by a defection from player 1, 2 expects himself to begin a punishment of, say T periods, and then to go back to cooperation forever. 2 expects 1 to play anything (i.e. neither D nor C will surprise him) in the next $T-1$ periods, then 2 expects 1 to go back to C (the ambiguity in 2's expectation is not necessary here, since 2 knows that if 1 follows the strategy, 1 will have $T-1$ periods of D to play for sure. The ambiguity will be useful after more complicated histories, so that when someone finishes punishing one neighbor, he can safely go back to C without further surprising the other neighbor). 2 also expects that 1 expects 2 to punish for T periods, and if 2 fails to punish, 2 will surprise 1 and trigger a more severe punishment. At the same time, 2 anticipates that the punishment will surprise 3 in turn, so he expects 3 to punish him for T periods. If 3 fails to fully carry out the punishment, then 2 expects himself to punish 3 for a much longer period of time, say bT periods, $b > 1$. Meanwhile, it is the mutual expectation of 2 and 3 that 3 should keep playing cooperation after 3 blocks. Finally 2 should also anticipate that his long punishment will keep surprising 1 later on, and 1 should react to it appropriately, and so on. A kind of social norm can thus be established by specifying people's expectations after any history. People then use the expectations to judge other people's behavior, and to guide their own.

Given any history of player 2, if he expects himself to play C in the next period, the strategy is going to be defined such that he also expects that at least one of the neighbors also expects him to play C, and if he deviates, he will postpone the time when the entire neighborhood goes back to C; if he expects himself to play D in the next period, he doesn't want to play C because it is either unnecessary to do so (when both 2 and 3 expect anything from him), or too costly to do so (when he expects at least one neighbor to expect him to play D). So far, sequential rationality is relative to the artificially defined expectation operators. Sequential equilibrium requires that the strategy be optimal with respect to real expectations formed under a consistent belief system. As we will see in later sections, essentially these real expectations are going to be duplicated by the artificial expectations, if we perturb the strategy approriately.

We can also see from this example why the mixed strategy doesn't work for finite number of players: 2 has incentive to mix between punishing and blocking only if 1 and 3 also have such incentive, but they don't.

1.3 Comparison with the Random Matching Model

The main message of this paper is this. The ultimate source of stability in this simple society is shared belief, or mutually compatible expectations on each other. An explicitly defined expectation system can be used as a social norm. What is important is not a common observation of a physical outcome, what is important is a common understanding of the social norm, the understanding that everybody knows the norm and is willing to follow it after any history.

This paper can be viewed as a special variation under a general theme, which is to disperse information in a repeated game among the players and ask whether the efficient outcome can still be maintained or not. There are other ways to disperse information. In Kandori(1992), people are pairwise matched at random in each period and play the prisoner's dilemma in each match. Each player only observes what goes on in his own matches, but not in other matches. Kandori constructs a contagion strategy that supports the cooperative outcome, provided that the cost of being defected upon is sufficiently large[4]. Bhaskar(1998) studied a simple transfer game in an overlapping generation setup. He found that with some mild informational constraints, transfers (from the young to the old in each period) cannot be supported by pure strategy equilibria. For example, if each generation only observes the actions of the past generation, then the only pure strategy equilibrium is global autarchy.

Apart from the superficial differences between our model and the random matching model, there are similarities as well as differences between the two. In both models, if l is large enough, the trigger strategy works for sufficiently large δ[5]; if l is small, the trigger strategy works for moderate values of δ. Moreover, Ellison's dilution idea applies to both models. The differences are more subtle: in the random matching model with public randomization, the supporting strategy is uniform with respect to δ, so long as δ is large enough. Without public randomization, however, the supporting strategy is not uniform with respect to δ. In addition, the equilibrium with public randomization is globally stable[6], but without public randomization, it is unknown whether global stability is possible. In our model, the strategy, call it E from now on, is both uniform with respect to δ and globally stable, for $l < 1$ and $g > 0$.

In the anonymous random matching model, a player's information about history can be summarized by a one dimensional statistic. It is impossible to keep track of other players' actions, and it is not necessary either. Although a player needs to guess how many players have been infected after any history, Kandori and Ellison simplified the analysis by constructing a strategy that is optimal against *any* reasonable guesses. Hence consistency is not an issue. In our model, however, the information is two dimensional, and a player has to treat them separately. Instead of implementing a T-period punishment scheme

[4]Ellison(1994) embedded a public randomizing device into the contagion strategy and showed that cooperation can be supported for any payoff parameters.

[5]In the random matching model the cutoff value of l depends on the population size, but it doesn't here.

[6]In the sense of Kandori(1992), global stability requires that players be able to return to efficient outcome eventually after any history.

probablistically, as in the strategy with public randomization in the random matching model, we carry it out deterministically here. The tradeoff is we have to specify a consistent belief system, because the strategy cannot be a best response for all belief systems.

This paper is organized as follows. Section 2 defines the pure strategy E, by defining a pair of expectation operators inductively. Section 3 shows that the strategy is sequentially rational with respect to the expectation operators. Section 4 derives a consistent belief system, under which real expectations can be formed after any private history. It is then shown that the real expectations formed after a history can be mimiced by the expectation operators after the same history. Section 5 discusses extensions to circles and trees. Section 6 concludes.