Available online at www.sciencedirect.com
ScienceDirect
Games and Economic Behavior 66 (2009) 785–812
GAMES and Economic Behavior
www.elsevier.com/locate/geb

Cooperative behavior and the frequency of social interaction

John Duffy *, Jack Ochs

Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260, USA

Received 30 June 2006

Available online 9 August 2008

Abstract

We report results from an experiment that examines play in an indefinitely repeated, two-player Prisoner's Dilemma game. Each experimental session involves N subjects and a sequence of indefinitely repeated games. The main treatment consists of whether agents are matched in fixed pairings or matched randomly in each indefinitely repeated game. Within the random matching treatment, we elicit player's strategies and beliefs or vary the information that players have about their opponents. Contrary to a theoretical possibility suggested by Kandori [1992. Social norms and community enforcement. Rev. Econ. Stud. 59, 63–80], a cooperative norm does not emerge in the treatments where players are matched randomly. On the other hand, in the fixed pairings treatment, the evidence suggests that a cooperative norm does emerge as players gain more experience.

JEL classification: C72; C73; C78; C92; D83

Keywords: Cooperation; Matching; Repeated Prisoner's Dilemma; Folk theorem; Community enforcement; Social norms; Experimental design

"Sometimes cooperation emerges where it is least expected."—Robert Axelrod, *The Evolution of Cooperation* (1984, p. 73).

1. Introduction

Cooperative behavior can expose individuals to possible exploitation by others who are willing to act opportunistically. Nevertheless, cooperation can be sustained if opportunistic behavior triggers a punishment that makes 'cheating' unattractive. Much cooperative behavior is sustained by decentralized, informal enforcement mechanisms. These mechanisms rely on individuals having an interest in how their current actions affect future social interactions. As Kandori (1992) notes, there are two general classes of informal mechanisms: Personal enforcement, where opportunistic behavior today destroys the possible benefits of future cooperation between the individual who has been cheated and the cheater; Community enforcement, where a cheater is sanctioned by other members of the community who have not themselves been victims of that cheater, but nevertheless refuse to engage in cooperative endeavors with any cheater. The Folk theorem holds for personal enforcement when a particular pair of agents has an indefinite

* Corresponding author at: University of Pittsburgh, Department of Economics, 4901 W. W. Posvar Hall, Pittsburgh, PA, USA. Fax: +1 412 648 1793.

E-mail address: jduffy@pitt.edu (J. Duffy).

0899-8256/$ – see front matter
doi:10.1016/j.geb.2008.07.003

number of future interactions with one another and the discount factor is sufficiently large. Kandori shows that public observability is sufficient for the Folk theorem to hold for community enforcement mechanisms as well. If the identity of the cheater is common knowledge then it does not make any difference if the cheater will have repeated future interactions with any particular member of the community or if the cheater will simply have an indefinite number of future interactions with various members of the community. That is, *regardless of matching protocol*, with public observability, there exist equilibrium strategy profiles that will support the same payoffs as are attainable under a fixed matching protocol.[1] More remarkably, Kandori shows that for groups of any fixed size there exist payoff functions such that there is a community enforcement mechanism that will sustain cooperation in an indefinitely repeated number of plays of a Prisoner's Dilemma stage game even when individual histories are purely private, individuals are anonymous, and the matching mechanism for each stage game is purely random.

In the anonymous, random matching case, it is the threat that one deviation from cooperation will trigger a contagious process of future defections by all who have experienced a defection—rather than the threat of being branded a 'cheater'—that acts as a deterrent. For this threat to be credible, an individual who has experienced a defection must find it more profitable to defect at the next opportunity, even though this will keep the contagious process going, than to continue to act cooperatively and stop the process. Of course, if only a small fraction of the population has already been infected, an individual may lose a considerable amount of future benefit from the eventual destruction of the cooperative norm. Therefore, for large groups, the threat of starting a contagious process will not be credible unless either the one time gain from cheating is very large, or there is some means of stopping the contagious process thereby limiting the loss from the destruction of the social norm. Ellison (1994) shows how the availability of a publicly observable randomization device may be used as a correlation device to signal the end of a punishment phase and resumption of a potentially new cooperative phase.[2]

While Kandori shows that a social norm of cooperation can be sustained as a non-cooperative equilibrium even when individual pairings are both random and anonymous, whether or not such norms are likely to emerge under such conditions is clearly an empirical question. The experiment described below was specifically designed to address that question.

Kandori's (1992) theorem applies to indefinitely repeated two-person games with minimal observation of the past actions of the individuals with whom a player is currently paired to play a stage game. Our experiment is designed, therefore, to study the behavior of individuals drawn from a fixed population who play an indefinite sequence of two-person Prisoner's Dilemma games under different matching protocols and different amounts of information transmission. The objective of varying the matching protocol (fixed pairings versus random matching) is to determine empirically how much difference in the level of cooperative play is associated with these different matching protocols. The objective of varying the information transmitted to players (under the random matching protocol only) is to determine whether information on the payoff or action history of a player's opponent, prior to play of the stage game, has any effect on the level of cooperative play. The design also incorporates a randomly generated, publicly observable signal that could be used by agents as a device to coordinate the end of a punishment phase, if they were to choose strategies of the type described by Ellison (1994).

To foreshadow our experimental results, we find that the initial play of subjects is quite similar under both fixed and random matching protocols. With experience, under fixed pairings, cooperation emerges as a norm. However, under random pairings, non-cooperation is quickly established as a norm, despite the presence of some conditional cooperators. Indeed, we provide direct evidence in one random pairings treatment of the presence of such conditional cooperators using a procedure introduced by Fischbacher and Gächter (2006) that builds upon the strategy method for

[1] Under public observability, an individual who has been labeled a 'cheater' has no incentive to act cooperatively in any particular interaction, as long as s/he believes that others will not cooperate with a known 'cheater.' Non-cheaters, in turn, will have little incentive to risk taking a cooperative action with a known 'cheater,' if they also believe that others will not cooperate with this particular individual in the future. Kandori provides two different ways of making individual deviations from cooperation unprofitable. One way depends upon individuals having group labels, independent of their actions. Individuals of one group are always matched with individuals of another group and once any member of a given group deviates, all members of that group are treated as 'cheaters,' by all members of the other group. Alternatively, individuals need not have group labels and only 'cheaters' are labeled. To prevent those who have not yet been labeled as a 'cheater' from finding it profitable to avoid acquiring this label, it must be the case that a non-cheater can expect to meet other non-cheaters sufficiently frequently in the future to make it profitable to give up the one time gain from cheating. This can be assured with strategies in which a cheater's label is removed after some finite number of stage games.

[2] As Ellison shows, it is not necessary to have a public randomization device to limit the period in which a contagion process is operating. But such a device does serve as a signal upon which all individuals can coordinate their departure from a punishment phase.