Journal of Behavioral and Experimental Economics 72 (2018) 1–8

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Economics

journal homepage: www.elsevier.com/locate/jbee

Mate choice mechanism for solving a quasi-dilemma

Tatsuyoshi Saijo[a], Junyi Shen[b,c,*]

[a] Research Center for Future Design, Kochi University of Technology, 2-22 Eikokuji, Kochi 780-8515, Japan
[b] Research Institute for Economics and Business Administration, Kobe University, 2-1 Rokkodai, Kobe, Hyogo 657-8501, Japan
[c] School of Economics, Shanghai University, 99 Shangda Road, BaoShan District, Shanghai 200444, China

ARTICLE INFO

Keywords:
Asymmetric prisoner's dilemma
Quasi-dilemma
Mate choice mechanism
Cooperation
Experiment

JEL codes:
C72
C92
D74
P43

ABSTRACT

Sajio et al. (2015) showed that the mate choice mechanism for a symmetric prisoner's dilemma (*PD*) game implements cooperation in backward elimination of weakly dominated strategies (*BEWDS*), and it attained almost full cooperation in their experiment. This study theoretically shows, first, that this mechanism works well in the class of quasi-dilemma (*QD*) games, such as asymmetric *PD* games and coordination games. Second, the class of *BEWDS*-implementable games is exactly the same as the class of *QD* games. Third, the mechanism cannot implement cooperation in a subgame perfect equilibrium. Finally, we confirm that the mate choice mechanism works well experimentally for an asymmetric *PD* game.

1. Introduction

Over the past six decades, the prisoner's dilemma (*PD*) game has been extensively discussed in both the public and academic press (see Poundstone, 2011 for a historical perspective). The *PD* game presents a scenario in which the outcome of one person's decision is determined by the simultaneous decisions of the other participants, resulting in a bad outcome for all of them (a Pareto-inefficient Nash equilibrium) if all act in their own self-interest. The key characteristic of this game is that while there are substantial gains that could be attained through cooperation, non-cooperation is dominant for each player (see Kuhn, 2014 for an overview of PD literature).

Since participants in laboratory experiments consider non-monetary factors such as social norms and anonymity as well as monetary stakes, the cooperation rates, that is, the ratio of participants who chose cooperation, in PD game experiments are well above zero (see Chaudhuri, 2011 for a survey), but not close to one. In order to increase the cooperation rates in PD games, researchers, such as Yamagishi (1986), Banks et al. (1988), Varian (1994), Fehr and Gächter (1999), Andreoni and Varian (1999), and Charness et al. (2007), started adding a stage *before* or *after* the PD games.[1] Adding a stage after the PD game allows participants to punish non-cooperation. Although participants in such games do not have a monetary incentive to punish other players, cooperation rates in experiments with punishment increase, but still do not become close to one (see Yamagishi (1986) and Fehr and Gächter (1999)). Varian (1994) introduced a compensation mechanism *before* the dilemma game. Under this mechanism each player is asked in the first stage to choose how much to pay his or her counterpart for cooperating. After learning the payments offered in the first stage, the players then play a normal *PD* game. Andreoni and Varian (1999) conducted experiments with the compensation mechanism and observed that cooperation rates increased from 25.8% when transfer payments were not feasible to 50.5% when transfer payments were permitted, and the cooperation rate was around 20% in the first two rounds. Charness (2007) conducted experiments and observed that cooperation rates increased from 11–18% when transfer payments were not feasible to 43–68% when transfer payments were permitted.

Our aim in this study is to find one of the *simplest* possible mechanisms to solve social dilemmas, including the prisoner's dilemma, both experimentally and theoretically. Experimentally, we focus on designing a mechanism that can attain a Pareto efficient outcome in a few rounds,[2] because we cannot repeat the same mechanism many

* Corresponding author.
E-mail address: shen@rieb.kobe-u.ac.jp (J. Shen).

[1] The choice of a participant in the first stage is scrutinized by the other participant before deciding the choice in the second stage. As Levitt and List (2007) suggested, this might increase cooperation.

[2] Chen (2005), for example, found that the stability property of mechanisms depends on their supermodularity. Supermodular mechanisms may require many periods to converge to a desired outcome. The goal of the endeavor is not to find such mechanisms but to find mechanisms that can attain a desired outcome in a few periods.

https://doi.org/10.1016/j.socec.2017.10.004
Received 22 February 2017; Received in revised form 20 October 2017; Accepted 23 October 2017
Available online 31 October 2017
2214-8043/

times in real-life settings. Theoretically, we do not stick to Nash or Nash-type equilibrium concepts, but search for a behavioral principle among subjects in experiments that implements cooperation. In addition, we do not use punishment or reward to balance the budget, that is, there is no monetary inflow or outflow in the mechanism design.[3]

The study by Saijo et al. (2015) is one of the first attempts to design such a mechanism for the *PD*.[4] They proposed the mate choice (*MC*) mechanism, which occurs *after* a *symmetric PD* game. After observing the choice of cooperation (*C*) or defection (*D*) in the *PD* game, each player is asked to approve or disapprove of the other's choice. If both approve it, the outcome is what they chose in the *PD* game; if at least one player disapproves of the other's choice, the outcome is the same as when both defect in the *PD* game.[5] Experimentally, Saijo et al. (2015) observed that the cooperation rate with the mechanism was 95.0% in round 1 and 96.9% over 19 rounds, when each subject was never matched with the same subject again in all rounds.[6] The (*C, C*) share, that is, the ratio of pairs in which both chose cooperation, was 90.0% in round 1 and 94.0% over 19 rounds. They also found that subjects' behavior was consistent with backward elimination of weakly dominated strategies (*BEWDS*) rather than Nash equilibrium (*NE*) or subgame perfect equilibrium (*SPE*) behavior. *BEWDS* is a procedure that eliminates weakly dominated strategies in each subgame, backwardly. The strategies that survive through the procedure are called *BEWDS* strategies. Theoretically, Saijo et al. (2015) proved that the *MC* mechanism implements cooperation in *BEWDS* for symmetric *PD* games.[7]

In two related studies, Masuda et al. (2014) constructed a minimum approval mechanism, which is a version of the MC mechanism, in a public good economy when the number of players is two and their preferences are linear. They showed experimentally that the mean contributions ranged from 76.9% to almost 100.0%, with an average of 94.9%. Huang et al. (2014) constructed a simplified approval mechanism, which is also based on the MC mechanism, for a symmetric *PD* game in which the number of players was expanded from two to three. They observed experimentally that the mean cooperation rate increased from 44.4% in round 1 to above 90.0% in round 5 and maintained that level in the remaining 10 rounds.

From the above-mentioned studies, it seems that the MC mechanism performs very well in stimulating the players to cooperate in a *symmetric* environment. Consequently, the question of whether the MC mechanism is also effective in an *asymmetric* environment is natural. As Andreoni and Varian (1999) found, a participant with a relatively low payoff tends not to cooperate and this might influence the cooperation rate. Hence, in this paper we expand the domain of this mechanism from *symmetric PD* games to *asymmetric* games that are not necessarily *PD* games. We find theoretically that the *MC* mechanism implements cooperation in *BEWDS* for the class of *quasi-dilemma* (*QD*) games, which contains coordination games, including the stag hunt game and *PD* games. Furthermore, under several mild conditions, we show that the class of games implementing cooperation in *BEWDS* is exactly the same as the class of *QD* games, and that the *MC* mechanism cannot implement cooperation in *SPE*.

In order to test the performance of the MC mechanism experimentally in an asymmetric environment, we choose an asymmetric parameterization of the *PD* game (``Game 3" in Charness et al. (2007)) because the cooperation rate of this parameterization (42.9%) was worse than those for the other two asymmetric parameterizations (53.9% in ``Game 1" and 68.1% in ``Game 2"). This fact also motivates us to investigate whether the MC mechanism performs better than the compensation mechanism does. It should be noted that the compensation mechanism does not cover all *PD* games; in contrast, the MC mechanism covers all *PD* games *and* non-*PD* games. That is, there is a class of *PD* games in which the compensation mechanism cannot implement cooperation in *SPE.* Game 3 belongs to this class.

Experimentally, we observed that the cooperation rate with the MC mechanism in an asymmetric environment started at about 76.7% in round 1, rose to 86.7% in round 2, 93.3% in rounds 3 and 4, 96.7% in round 5, and then stayed above 98.0% in the remaining 14 rounds. The overall average cooperation rate over 19 rounds was 96.7%. The (*C,C*) share started at 56.7% in round 1, rose to 73.3% in round 2, to 86.7% in rounds 3 and 4, to 93.3% in round 5, and then stayed above 96.0% in the remaining 14 rounds. The overall average (*C,C*) share over 19 rounds was 93.5%. That is, the MC mechanism works reasonably well, although it took a few rounds to achieve high (*C,C*) share in an asymmetric *PD* game.

There are many examples of the mate choice mechanism. Consider a merger or a joint project of two companies. They must propose plans (the contents of cooperation) in the first stage, and then each faces the approval decision in the second stage. In order to resolve the conflicts such as prisoner's dilemma, interested parties usually form a committee consisting of representatives of the parties. Consider two companies facing confrontation on the standardizations of some product. Each company chooses cooperation (or compromise) or defection (or advocating of the own standard), and then the committee consisting of two company members and/or bureaucrats gives the approval. Another example is the two party system. Each party chooses either cooperation (or compromise) or defection (or insistence of policy for the own party), and then diet (or national assembly) plays a role of approval. The bicameral system also has two stages. One chamber decides a policy (or compromise) and the other chamber plays a role of approval. The negotiation process at United Nations also has this structure. Negotiators among relevant countries get together to find compromise, i.e., the content of cooperation in the first week and then high ranked officials such as presidents and prime ministers get together to approve or disapprove it in the second week. Adding the second stage in resolving conflicts has been used widely in our societies.

The paper is organized as follows. Section 2 describes the MC mechanism applied to *QD* games. Section 3 proves that *BEWDS* implementable games are *QD* games and shows that the *MC* mechanism cannot implement cooperation in *SPE*. Section 4 presents the experimental design, and Section 5 presents the results. Section 6 provides suggestions for further research.

2. The mate choice mechanism and quasi-dilemma games

Consider a 2×2 game that has two strategies: cooperation (*C*) and defection (*D*).

Define a payoff function p for the game as follows: $p(C,C) = (a,v) = V$, $p(D,C) = (c,x) = X$, $p(C,D) = (b,w) = W$, and $p(D,D) = (d,z) = Z$. If p satisfies $V > Z$ ($a > d$ and $v > z$), $X \neg\geq Z$ ($d > c$

Player 2

		C	D
Player 1	C	(a,v)=V	(b,w)=W
	D	(c,x)=X	(d,z)=Z

Fig. 1. A *QD* game.

[3] According to Guala (2013), *strong reciprocity*, in which a player punishes other players using the player's own resources, is rare in human history.

[4] Saijo et al. (2016) is a simplified version of Saijo et al. (2015).

[5] Because the *MC* mechanism does not have devices such as punishment or reward, it is budget balanced.

[6] This is called complete stranger matching, and only a few experiments employ this matching. Saijo et al. (2015) chose this matching since it is the least favorable design for cooperation with respect to matching.

[7] The *MC* mechanism uses unanimity. Bankset al. (1988) introduced a voting stage after a public good provision stage and observed that unanimity reduced efficiency. Researchers stopped pursuing this avenue after Banks et al. (1988) presented their findings. Furthermore, Masuda et al. (2014) found that the *MC* mechanism cannot implement a Pareto-efficient allocation in *BEWDS* for an economy with a public good.