Economics Letters 124 (2014) 199–202

Contents lists available at ScienceDirect

Economics Letters

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

Conditional punishment

Kenju Kamei *

Department of Economics, Bowling Green State University, United States

HIGHLIGHTS

- We elicit human conditional punishment types by conducting experiments.
- The basic game frame used in this study is the voluntary contribution mechanism.
- A strategy method is used to elicit individual conditional punishment types.
- The more others punish a person, the stronger people on average punish him.
- There exist heterogeneous individual conditional punishment types.

ARTICLE INFO

Article history:
Received 27 February 2014
Received in revised form
12 May 2014
Accepted 21 May 2014
Available online 29 May 2014

JEL classification:
C91
C92
D70
H41

Keywords:
Cooperation
Punishment
Experiments

ABSTRACT

We elicit human conditional punishment types by conducting experiments. We find that their punishment decisions to an individual are on average significantly positively proportional to other members' punishment decisions to that individual.

1. Introduction

Cooperation problems are one of the most common and important features of our modern society. During the recent few decades, economists have devoted efforts to understanding people's cooperation behavior in dilemma situations and have found that peer-to-peer punishment opportunities can be effective in promoting cooperation under some conditions (e.g., Fehr and Gächter, 2000, Nikiforakis and Normann, 2008). They have also demonstrated that neither people's cooperation decisions nor punishment decisions are homogeneous. It is now known that there are multiple human types with regard to cooperation and punishment. Regarding cooperative types, it has been known that some people have preferences for cooperation in dilemmas. Moreover, many researchers have proposed that there exist some people that conditionally cooperate dependent upon decisions of others (e.g., Fischbacher et al., 2001, Kurzban and Houser, 2005). As for punishment types, there are many studies finding not only various unconditional punishing behaviors, but also counter-punishing behaviors. For instance, anti-social punishers impose a fine on cooperators (e.g., Cinyabuguma et al., 2006, Herrmann et al., 2008). Moreover, when some low cooperators are punished by high cooperators, they anti-socially retaliate against the high cooperators, given a counter-punishing opportunity (e.g., Nikiforakis, 2008). Some people engage in punishment of non-punishers or that of anti-social punishers (e.g., Denant-Boemont et al., 2007). However, surprisingly little attention has been paid to how people's punishment decisions to a person may depend on punishment decisions of others to that person.

Studying people's conditional punishment behaviors is important especially for three reasons. First, understanding a person's punishment decisions in relation to others' contributes to an advancement in the economic theory as it gives us more

* Tel.: +1 419 372 6868.
E-mail addresses: kenju.kamei@gmail.com, kkamei@bgsu.edu.

http://dx.doi.org/10.1016/j.econlet.2014.05.017
0165-1765/

micro-level data regarding people's punishment behaviors. Second, if multiple types of punishers are present, then, incorporating the task of eliciting the punisher types in experiments, likewise cooperative types elicited in other past studies, would allow us to interpret subjects' decisions more easily and precisely (e.g., by studying the relation between conditional punishment types and their institutional choices). Third, providing evidence on the prevalence of heterogeneous conditional punishment types would stimulate both theoretical and experimental research on the roles of conditional punishment types in the evolution of cooperation norms. For instance, Szolnoki and Perc (2013) theoretically show that what they call conditional punishers – those who impose a fine with a strength proportional to the number of punishers in their own groups – can play an important role in promoting cooperation.[1]

This paper elicits people's conditional punishment types by conducting experiments using a strategy method. We find that subjects on average punish a non-cooperator at a strength positively proportional to the punishment by other members. We find, however, that subjects' individual conditional punishment decisions are heterogeneous and that around half of the subjects do not punish a person, no matter how many punishment points other members assign to that person. We call this punishment type the free-rider. Moreover, our data indicates that the classification of conditional punishment types helps us explain people's unconditional altruistic as well as anti-social punishment decisions.

The rest of the paper proceeds as follows: Section 2 describes our experimental design. Section 3 reports results, and Section 4 concludes.

2. Experimental design

The basic game frame we use in the experiment is the voluntary contribution mechanism. Each subject is randomly assigned to a group of four, is given an endowment of 10 points, and decides whether they allocate the 10 points to either their private account or public account. When a subject allocates it to her private account, she will obtain 10 points as her payoff. By contrast, when the subject allocates it to her group's public account, all group members, including her, obtain 5 points each as their payoffs. In other words, the MPCR (Marginal Per-Capita Return) is 0.5. We call those who allocate their endowment to their public (private) account the cooperators (non-cooperators) in this paper.

Once all subjects make their allocation decisions, they subsequently make conditional punishment decisions. Each subject can assign up to four punishment points to each member. The punishment points must be an integer. For each punishment point he assigns to a member, he loses one point from his payoff, and the punished loses three points from her payoff. In the conditional punishment stage, each subject makes 18 ($= 2 \times 9$) decisions, contingent on whether the target is a cooperator or a non-cooperator (2 possibilities), and also on how many punishment points on average the other two members in his group assign to the target ($= \{0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0\}$), of which there are nine possibilities. We refer to this list of 18 punishment decisions as the "conditional punishment schedule". Once all subjects complete their conditional punishment schedules, they move on to the unconditional punishment stage. In this stage, they are informed of the allocation decisions made by the three other members in their groups and then make unconditional punishment decisions to each member. The conditional punishment decision is incentive-compatible as follows: once everyone makes their unconditional punishment decisions, one schedule out of the four in each group is randomly selected to be used as the conditional punishment schedule.[2] The instructions are available online as a supplementary content (see Appendix A).

Standard theory predictions are straightforward: all allocate their endowments to their private accounts, and no one penalizes others. However, social preference models such as Fehr and Schmidt (1999) predict that some allocate their endowments to the public account and some punish non-cooperators. Assuming cooperator j is only concerned about the income inequality with non-cooperator i, the income inequality-averse model predicts more punishment from j to i when the other two members punish i less, because otherwise the income inequality between i and j would be large (see Appendix Table B.1).

3. Results

We conducted experiments at University of Michigan–Ann Arbor in February 2014. All experiments were neutrally framed. 52 undergraduate students participated in the experiments. No subject participated in more than one session. The experiments lasted around 45 min on average. The average per-subject payment (including a participation fee of $5) was $17.35.

Fig. 1 reports the total average conditional punishment schedule in our experiment.[3] The total average conditional punishment decisions to a non-cooperator are significantly increasing in the amount of the other members' average punishment to the non-cooperator. This pattern cannot be rationalized if we assume that a decision-maker is only concerned about the income inequality between her and the punished (see predictions based on inequality-averse models in Section 2). This instead suggests that (a) she does not care how small a payoff the non-cooperator obtains but (b) she does care about the income inequality between her and other punishers.

Moreover, intriguingly, we find that the total average conditional punishment decisions to a cooperator are also significantly increasing in the others' punishment toward the cooperator, although the strength of punishment is significantly less than that to a non-cooperator. This pattern cannot be explained by reciprocity or income inequality-averse models since the cooperator behaves nicely to his group members and his payoff cannot be any higher than theirs. This pattern can be, however, explained by Levine (1998) which assumes that some individuals have spiteful preferences.

Result 1. *The total average conditional punishment decisions are positively proportional to the other members' punishment decisions, whether the target of punishment is a cooperator or non-cooperator.*

Fig. 1 also shows the average conditional punishment schedules for each of the cooperators and non-cooperators. The total average punishment schedules of both cooperative types resemble the one of all subjects, but the strength of punishment differs by cooperative type. We find that on average non-cooperators conditionally punish both cooperative types at a statistically similar strength, whereas cooperators conditionally punish a non-cooperator more strongly.[4]

Result 2. *Non-cooperators on average conditionally punish both cooperative types at a statistically similar strength, but cooperators on average conditionally assign significantly larger punishment points to a non-cooperator than to a cooperator.*

[1] See also Boyd et al. (2010).

[2] This procedure is often used in similar studies including Fischbacher et al..

[3] See Appendix Table B.2(1) for regression results.

[4] See Appendix Table B.2(2).