Economics Letters 171 (2018) 193–197

Contents lists available at ScienceDirect

Economics Letters

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

The role of visibility on third party punishment actions for the enforcement of social norms

Kenju Kamei

Department of Economics and Finance, Durham University, United Kingdom

HIGHLIGHTS

- A prisoner's dilemma game experiment with a third party punisher was conducted.
- Third parties do not punish cooperators when their punitive actions are visible.
- Third parties impose strong punishment on a norm violator with high visibility.
- With high visibility, a player receives a higher expected payoff by cooperating.

ARTICLE INFO

Article history:
Received 3 March 2018
Received in revised form 7 June 2018
Accepted 30 July 2018
Available online xxxx

Keywords:
Experiment
Cooperation
Dilemma
Third party punishment
Social norms

JEL classification:
C92
D01
H49

ABSTRACT

This paper presents results from a prisoner's dilemma game experiment with a third party punisher. Third party punishment was frequently observed, in line with previous studies. Despite the prevalence of punishment, having one third party punisher in a group did not make one's defection materially unbeneficial because of the weak punishment intensity observed. When a third party player's action choice was made known to another third party player in a different group, however, third party punishment was sufficiently strong to transform the dilemma's incentive structure into a coordination game, through which cooperation norms can be effectively enforced.

1. Introduction

One well-known and consistent finding in recent decades is that some people display other-regarding preferences, such as inequity aversion, when interacting with others. A large body of experimental research has shown that even third parties, who are not directly involved in the relevant interactions, frequently impose punishment when they encounter unfair economic behavior in dilemma games (e.g., Fehr and Fischbacher, 2004; Lergetporer et al., 2014; Kamei, forthcoming)

Most research in this area to date found that while third party punishment is frequently observed, it is much weaker than direct punishment (e.g., Fehr and Fischbacher, 2004). This paper experimentally studies how the visibility of third parties' punitive actions may affect their punishment behaviors. This research question is motivated by past work proposing that the visibility of actions enhances people's pro-social behavior through image motivation (see, e.g., Bénabou and Tirole (2006) for a theoretical model, and Ariely et al. (2009) for experimental evidence on charitable giving). It is also motivated by the research which suggests that increasing the visibility of actions within a group may affect people's altruistic tendencies in the ongoing interactions (see, e.g., Sell and Wilson (1991) for the impact of individualized, instead of aggregate, information on voluntary contributions to public goods, and Kamei and Putterman (2015) for direct higher-order punishment in a public goods game). High visibility of action choices may trigger social effects, such as shame and pride (e.g., Bowles and Gintis, 2005), potentially influencing third parties' punishment behaviors.

In the experiment, there are two players that engage in a one-shot prisoner's dilemma game with each other (PD players, hereafter), and a third party player who decides how to impose sanctions on the PD players in each group. The results demonstrate that the punishment intensity on a norm violator is much stronger when each punisher's action choice is made known to another punisher, than when the punitive actions are kept private. Moreover, in the high visibility condition, the third party players almost completely refrain from perverse punishment of cooperators. These

E-mail addresses: kenju.kamei@durham.ac.uk, kenju.kamei@gmail.com.

https://doi.org/10.1016/j.econlet.2018.07.043
0165-1765/

findings suggest that raising the visibility of third parties' punitive actions can be a powerful device for disciplining their sanctioning activities.

2. Experimental design and hypotheses

This study is based on a prisoner's dilemma game with a third party player (Fehr and Fischbacher, 2004). There are two treatments, namely the "Standard" and "Visibility" treatments, implemented using a between-subjects design. Experimental points are converted into pounds sterling at a rate of five points to £1.

2.1. The standard treatment

At the onset of this treatment, subjects are randomly assigned to a group of three so that each group has two PD players and one third party player. There are two stages. In Stage 1, PD players are each endowed with 25 points and simultaneously decide whether to send 10 points to their counterparts. If a subject sends 10 points to her counterpart, the counterpart receives 30 (= 3 × 10) points and the remaining 15 points become the sender's payoff. If the subject does not send 10 points, she retains the full endowment as her payoff. Hereafter, we call a subject who sends (does not send) 10 points a "cooperator" ("defector"). The third party player in each group is not involved in the prisoner's dilemma interaction.

In Stage 2, each third party player receives an endowment of 40 points and makes punishment decisions for their respective group. Punishment points assigned to each PD player must be an integer between 0 and 20. For each punishment point assigned to a target, one point is deducted from the third party player's payoff and three points are deducted from the target's payoff. Each punisher makes the following four decisions using the strategy method (punishers make decisions before being informed of the first stage outcome):

Scenario CC**:** Punishment points targeted at a cooperator who interacted with another cooperator;

Scenario DC**:** Punishment points targeted at a defector who interacted with a cooperator;

Scenario CD**:** Punishment points targeted at a cooperator who interacted with a defector;

Scenario DD**:** Punishment points targeted at a defector who interacted with another defector.

After third party players complete four decisions, their choices corresponding to the realized PD players' sending decisions are applied.

Standard theory predicts no punitive behaviors of third parties because punishment is privately costly. As shown in Appendix C.1, however, the inequity aversion model of Fehr and Schmidt (1999) suggests that (a) a third party player i punishes a defector in Scenario DC if i exhibits sufficiently strong aversion to disadvantageous inequality (i.e., $\alpha_i > 1 - \frac{\beta_i}{2}$), and (b) i even punishes a cooperator in Scenario CC if i exhibits much stronger aversion to disadvantageous inequality (i.e., $\alpha_i > 2$).

2.2. The visibility treatment

The Visibility treatment is identical to the Standard treatment, except that each third party player is randomly and anonymously paired with another punisher in a different group, akin to an enforcement team, and their respective punishment behavior is made known to the partner. This visibility condition is common knowledge to all subjects. Even though two third parties are put in a team, they act independently to make punishment decisions toward different PD players. There is real-world relevance here: for instance, individuals who work in public enforcement usually share reports with other officers working in the same role in the event of encountering law violators.

If i has (non-strategic) image motivation (e.g., Ariely et al., 2009; Bénabou and Tirole, 2006), i may punish a cooperator less in Scenario CC and punish a defector more in Scenario DC in the Visibility than in the Standard treatment. As we can reasonably assume that a non-trivial fraction of subjects are concerned about their image, we can formulate the following hypotheses in the paper:

Hypothesis 1. Punishment strength in Scenario CC is weaker in the Visibility than in the Standard treatment.

Hypothesis 2. Punishment strength in Scenario DC is stronger in the Visibility than in the Standard treatment.

The Fehr–Schmidt model also predicts Hypothesis 1, because as illustrated in Appendix C.2, inequality averse i's punishment strength in Scenario CC would be positively correlated with i's beliefs on her matched punisher j's punishment strength in this scenario. This implies that i may refrain from engaging in punishment of cooperators in Scenario CC, considering that usually only a minority of subjects commit such anti-social punishment (thus i would form a belief that her counterpart is less likely to punish cooperators). The model does not, however, predict Hypothesis 2. This is because (a) i's disutility resulting from inequality with someone in j's group increases if i attempts to match her punishment with j's strength in Scenario DC (note that $\alpha \geq \beta$), and (b) when i makes punishment decisions in Scenario DC, j does not necessarily confront with Scenario DC(CD) — he may also confront with Scenario CC or DD with some probability (see the Appendix for the detail).

In the experiment, the identities of all subjects are not disclosed in order to measure the pure impact of high visibility on third parties' punishment behaviors.

2.3. Experimental procedure

Four sessions were conducted at the EXEC laboratory at the University of York in December 2015 and February 2016. A total of 96 students (48 students per treatment) participated in the experiment. No subjects participated in more than one session. All experiments except instructions were programmed using the z-Tree software (Fischbacher, 2007). The instructions and verbal explanations in the experiment were neutrally framed.

3. Results

PD players' cooperation rates were the same for the two treatments at 71.9% (23 out of 32 PD players). This implies that PD players did not expect changes in visibility would affect third parties' behaviors. Third parties' punishment behaviors, however, were very different between the treatments.

3.1. Punishment decisions of third party players

The pattern of punishment replicates that of past research. Fig. 1 indicates that third party punishment is common, and that both its frequency and strength are much higher in Scenario DC than in any other scenario.[1] This pattern resonates with the idea that people are inequality-averse and that third party players attempt to mitigate income inequality by inflicting punishment.

[1] The frequency of punishment in Scenario DC is significantly higher than that in Scenarios CC and CD in the Visibility treatment (Appendix Table A.1). The punishment intensity in Scenario DC is significantly stronger than that in Scenarios CC and CD in both the treatments (Appendix Table A.2).