Games and Economic Behavior 82 (2013) 91–102

Contents lists available at SciVerse ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Punishment strategies in repeated games: Evidence from experimental markets

Julian Wright [1]

Department of Economics, National University of Singapore, Singapore

ARTICLE INFO

Article history:
Received 6 December 2011
Available online 9 July 2013

JEL classification:
C72
C73
C90
D43
D83

Keywords:
Experiment
Cooperation
Tit-for-tat
Grim trigger strategy

ABSTRACT

An experiment is designed to provide a snapshot of the strategies used by players in a repeated price competition game with a random continuation rule. One hundred pairs of subjects played the game over the Internet, with subjects having a few days to make their decisions in each round. Occasionally subjects are asked to enter one-period-ahead pricing strategies instead of prices. According to the elicited strategies, between 90% and 95% of subjects punish less harshly (in their initial response to a deviation) than implied by the grim trigger strategy, and do so in a way that depends on the size of the other subject's deviation. Future earnings are highest for subjects adopting the tit-for-tat strategy, even after controlling for a subject's past earnings. Punishment strategies are generally softer and more graduated than implied by a grim trigger strategy, and do better as a result.

1. Introduction

When decision makers interact repeatedly through time, we know from previous lab experiments that cooperative outcomes can arise in competitive settings provided there are not too many competitors (Engel, 2007). However, we know far less about the strategies used to achieve these outcomes. In theories and applications of infinitely repeated games, economists have focused almost exclusively on certain equilibrium trigger strategies, which I will refer to as *disproportionate punishment strategies*. These strategies have three important features in case of any player defecting from the cooperative agreement — the punishment is immediate, the punishment is harsh (often maximal), and the same harsh punishment applies regardless of the nature of the deviation (i.e. it is independent of the "crime"). The prime examples of disproportionate punishment strategies are: (i) grim trigger strategies (hereafter, Grim) following Friedman (1971) in which agents cooperate initially but revert to the one-shot Nash equilibrium forever immediately following any agent defecting from the cooperative agreement, and (ii) optimal symmetric two-phase punishment strategies (hereafter, Optimal) following Abreu (1986, 1988) which are similar except that the punishment phase can be even harsher than playing one-shot Nash but it is only played for a fixed number of periods, after which players return to the cooperative phase unless there is a deviation from the punishment phase in which case the punishment phase is restarted.

I make use of a restricted version of Selten's (1967) strategy method to address whether experimental evidence provides any support for these types of disproportionate punishment strategies, focusing on three important features of punishment

E-mail address: jwright@nus.edu.sg.

[1] I thank Simon Anderson, Ananish Chaudhuri, Chew Soo Hong, Charlie Holt, Nathan Larson, Maroš Servátka, Satoru Takahashi, an editor and three referees for their very helpful comments, as well as Yang Jun, Derrick Kam, Xiang Li and Huan Yuen Tan for their excellent research assistance. I also gratefully acknowledge research funding from the Ministry of Education AcRF Tier 1 fund under Grant No. R122000105-112.

0899-8256/$ – see front matter
http://dx.doi.org/10.1016/j.geb.2013.06.012

implied by these strategies (they are immediate, harsh, and independent of the "crime"), or whether it provides support for any other types of strategies.

In the experiment, 200 university students are randomly assigned into 100 duopoly markets. Each subject faces a payoff table which comes from a symmetric Bertrand duopoly with differentiated products and is the same for all markets. Subjects are paired with one other subject throughout the experiment. The setting is one of perfect monitoring with a random continuation rule. After an introductory lab session, subjects set their prices through a website. The experiment is designed so that subjects have plenty of time to reflect on their choices (2–3 days per round compared to 1–2 minutes in a lab session). As a result, some markets lasted for over six months.

While normally each subject in a market is asked for their price after observing the other subject's previous price, occasionally (as determined by a random draw) subjects in a particular market and particular round do not see each other's previous price. Instead, subjects are asked for their one-period-ahead pricing strategy (the price they wish to set for each possible price the other subject might have set in the previous round). Since their response is only elicited for one period ahead, this is a restricted form of Selten's "strategy method". A subject's given one-period-ahead strategy, together with the price the other subject actually set in the previous round, determines their price for the round. By only asking for strategies occasionally, subjects first experience setting prices in the normal fashion. This design allows me to establish that asking for strategies rather than prices does not distort decisions, thereby validating the use of this method for the present experiment. The design also minimizes the number of contingent choices faced by subjects, which is particularly important given the action set is large, and does so without restricting a subject's choice in any way as would be the case if subjects had to specify such strategies from the start. For expositional convenience, the elicited one-period-ahead strategies (equivalently, intertemporal response functions or conditional actions) will be referred to as "strategies" throughout the paper.

The evidence from the experiment suggests very few subjects adopt disproportionate punishment strategies. Only about 5% of subjects have elicited strategies that are consistent with them cooperating using these types of strategies. Instead, subjects tend to use graduated punishments like tit-for-tat, with prices that vary with the other subject's previous price, if they indeed intend to use any immediate punishment at all.[2] Between 90% and 95% of subjects have elicited strategies that are less harsh in their immediate response to a negative deviation from the cooperative price than implied by Grim, with more than 20% of subjects having elicited strategies that are lenient (they do not involve any immediate response to such undercutting).

The adoption of strategies other than disproportionate punishment strategies does not seem to come at any cost to the subjects involved. Based on regressions of future earnings on strategy choices, subjects that adopt disproportionate punishment strategies do not enjoy significantly higher present discounted value of earnings compared to the tit-for-tat or lenient strategies. Indeed, subjects that adopt the tit-for-tat strategy enjoy significantly higher future earnings, even after controlling for past earnings, with the present discounted value of earnings estimated to range from US$14 to US$24 more than for disproportionate punishment strategies, depending on the specification adopted.

The present paper relates to recent experimental studies that draw inferences on strategies used in repeated game settings by observed outcomes. Most such studies use repeated prisoner's dilemma games. Recent contributions find evidence against Grim (Camera et al., 2012) and in favor of the use of tit-for-tat strategies (Dal Bó and Fréchette, 2011), or variations of tit-for-tat strategies in which subjects are initially lenient, waiting to see if any deviation persists before reacting with punishment (Fudenberg et al., 2012).[3]

In contrast to these studies, I directly recover a snapshot of subjects' strategies by recovering their one-period-ahead strategies in occasional rounds. In a recent working paper, Dal Bó and Fréchette (2012) also elicit strategies directly.[4] In their treatment with a 0.9 continuation rate (i.e. the closest to the 0.95 rate I use), tit-for-tat is the most popular strategy and its popularity increases as the subjects gain experience. An important difference between my paper and this and other works based on the repeated prisoner's dilemma game is that I consider a competition game where the action space is richer, so the extent of punishment can vary with the extent of the "crime". In addition to whether subjects use punishment immediately following a deviation, this allows me to explore how harsh any such punishment is, and whether it depends on the extent of the rival's deviation. In the prisoner's dilemma game, subjects have only two choices (cooperate and defect) so that these distinctions are not possible. Thus, while Fudenberg et al. (2012) focus on the time dimension of punishment (how long it takes for punishment to come into effect following a defection, for a fixed punishment level), I focus on the punishment immediately following a defection but allow subjects a full range of choices to determine how harsh their initial punishment is.

In allowing for a richer action set than the two choices available each round in prisoner's dilemma games, the closest work is Selten et al. (1997). They have each subject supply a strategy which is played against every other subject's strategy

[2] A small literature (Kalai and Stanford, 1985; Samuelson, 1987; Friedman and Samuelson, 1990, 1994; and Lu and Wright, 2010) has studied cooperation in repeated games with graduated punishments. However, except for the work of Slade (1992) and more recently Garrod (2012), these types of punishment strategies have not been employed in the applied literature.

[3] Mason and Phillips (2002) test between the implications of different types of trigger strategies, providing evidence consistent with longer-lived, less intense punishment phases than short-lived but intense punishment phases. Engle-Warnick and Slonim (2006) infer strategies from an infinitely repeated trust game using deterministic finite automata to back out the strategies that best fit the observed actions.

[4] They consider a repeated prisoner's dilemma game in which subjects specify full strategies from the start, but to reduce the number of contingencies, they give subjects a limited menu of strategy choices.