Punishment strategies in repeated games: Evidence from experimental markets

Julian Wright*

May 2013

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

*Department of Economics. National University of Singapore: email jwright@nus.edu.sg. 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.

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 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 Selton'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.[1] Between 90% to 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.*, 2011).[2]

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.[3] In their treatment

[1] 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.

[2] 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.

[3] They consider a repeated prisoner's dilemma game in which subjects specify full strategies from the

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.* (2011) 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 (one-by-one) in a series of 20-period asymmetric Cournot duopoly supergames through a computerized tournament, with the goal being to supply a strategy which does the best against all others. They find subjects achieve cooperation best by a "measure-for-measure policy," which reciprocates movements towards and away from the ideal point by similar movements. Thus, they provide a generalization of the finding of Axelrod and Hamilton (1981) that tit-for-tat does best in computerized tournaments of prisoner's dilemma games.

Compared to Selten *et al.*, my setting is more standard in that subjects are matched into fixed symmetric duopoly markets rather than in tournaments, with subjects paid depending on their earnings. I am also less ambitious in the information elicited. Selten *et al.* obtain subjects' full strategies (for all possible histories) by having subjects program their strategies in advance on a computer. Given the complexity of defining some history-dependent strategies (e.g. for a duopoly with N price choices, there are N^{2T} possible histories after T rounds), there is the real risk that requiring full strategies be described from the start could distort or artificially simplify subjects' chosen strategies. In comparison, I only occasionally elicit subjects' one-period-ahead strategies during the course of the experiment, which is a lot more tractable for subjects but still allows them to choose their desired action in every round based on the full history of play to that point. Despite these differences, I also find the tit-for-tat strategy does best.

A literature on public goods games (e.g. Fischbacher *et al.*, 2001) has also used the strategy method to elicit strategies in a setting where subjects have access to a wide

start, but to reduce the number of contingencies, they give subjects a limited menu of strategy choices.

range of possible actions (their contribution levels). However, the strategies elicited in Fischbacher et al. (2001) correspond to each member's best-response to the average contribution level of other group members in a one-shot game, so zero contribution is a dominant strategy for each member. This is in contrast to the one-period *intertemporal* best-responses elicited in the infinitely repeated game setting I consider, in which cooperation can be rational for self-interested subjects. This raises the question of whether the popularity of graduated punishment strategies reflects a non-strategic preference by subjects for reciprocity. While I cannot rule out this possibility, the market framing I employ, in which each subject is identified as a seller that sets prices should make this less likely. Indeed, Reuben and Seutens (2012) provide some recent evidence using the strategy method that suggests that strategic concerns are the primary motivation for cooperation in repeated prisoner's dilemma games.[4]

The rest of the paper proceeds as follows. Section 2 outlines the experimental design. Section 3 outlines the different types of theoretical strategies I will focus on. Section 4 details the results from the experiment, while Section 5 concludes.

2 Experimental design

A Supplementary Appendix contains the full experimental design, instructions, payoff table, and some screen shots. Here I focus on the key aspects of the design.

The experiment was conducted at the National University of Singapore and through the Internet from August 2007 to March 2008. Students were recruited from 50 different classes, and enrolled subjects attended one of 13 separate briefing sessions. The briefing sessions covered the instructions (which were read out aloud) as well as trial rounds and a short test. The 200 subjects were randomly paired into 100 markets such that no two subjects from the same major were matched. This design made it unlikely matched subjects would know each other given a student population of around 29,000.

All subjects received the same payoff table which showed their payoff (in lab dollars L$) for each combination of their own price and the other subject's price. The payoffs are derived from a symmetric Bertrand duopoly game with differentiated products. Prices are the whole numbers (lab dollars) from 1 to 25. The unique one-shot Nash equilibrium price p^n is L$2. The corresponding payoff is denoted π^n. The jointly optimal cooperative price (i.e. the monopoly price) p^m is L$23. In addition to the S$10 for participating in the lab session, subjects received S$1 for every 50 lab dollars they obtained. (S$10 $\simeq$ US$7 at the time). Their payoff was L$45.02 or S$0.90 per round at the one-shot Nash equilibrium price, L$270.65 or S$5.41 per round at the monopoly price, and L$525 or S$10.50 in any

[4] See also Cabral *et al.* (2012) and Dreber *et al.* (2011) for other related evidence.

round in which they chose L$21 while the other subject chose the monopoly price (i.e. they chose the myopic best response to the monopoly price). Subjects could enter their prices via the Internet at any time up to a fixed cutoff time for each round, which was at 8pm on a Tuesday, Thursday and Sunday of each week.

In the instructions, subjects were told that they represent a “seller” of a product or service and that they have to decide what price to set. They were matched throughout with one other subject, known to them only as the “other seller”. In the first round and most subsequent rounds, subjects had to decide which price to set (from L$1 to L$25). Subjects could see their total payoff in lab dollars and could click on a link to see past choices (theirs and the other seller’s) and their associated payoffs.

Subjects were informed that after round 12, there was a 5% chance each round that a market would end, upon which both subjects would be paid their cumulative earnings. It was also possible that a subject did not enter their price by the cutoff time, despite reminders. In this case, as indicated in the instructions, the affected market was closed, the offending subject received no payment, and the other subject was paid in a way which neither particularly rewarded or punished him for this event, as had been stated in the briefing session. This avoids the problem, that a subject may exit the experiment in order to impose a more severe punishment on the other subject, albeit at a personal cost (which Camera and Casari, 2009, show can have important implications for cooperation).

Subjects were also informed that sometimes a round may differ in that instead of entering a single price, they would have to enter a pricing strategy. This, together with the price the other subject actually set in the previous round, determined their actual price for the round. Unlike prices, a subject’s past strategy is not observed by the other subject they are matched to (only the price that results from it is observed subsequently).

In theory, it would make no difference to a subject whether a round is a strategy round or a price setting round. The available choices and associated payoffs are essentially the same. However, asking for strategies in every round would likely cause confusion among subjects since they would never see the actual prices set by their opponent in the preceding round.[5] Instead, strategies were only asked starting from round 10, after subjects had first experienced setting prices. Specifically, in rounds 10, 15, 20, ... each market which was still open had a 25% chance of being selected as a strategy round if the market was not going to end within the subsequent five rounds and a 75% chance of being selected as a strategy round if the market was going to end within the subsequent five rounds. To implement this approach, the length of each market was fixed prior to the beginning of the experiment based on a simulation using the 5% chance each market will end each

[5]This was the main feedback from a pilot experiment that was run earlier with a group of 88 business and economics students for 20 rounds, in which one-period-ahead strategies were elicited in every round.

round (after round 12) and entered into the computer program that ran the experiment over the Internet. This design ensured that subjects would not get asked for strategies too many times, especially during the earlier part of their market's existence, while at the same time increasing the chance of eliciting their strategies once their cooperative behavior had stabilized. Subjects were not informed of this particular formula; in the lab briefing they were only told that "sometimes" a round will be a strategy round. By only asking for strategies in rounds that are multiples of five, and by keeping randomness in the decision to ask for strategies, subjects would be unlikely to be able to work out any linkage between being asked for a strategy and the likelihood their market would end.

The types of strategies subjects use to sustain cooperative outcomes could be different if subjects can communicate their intentions to each other. To explore the possible effect of communication on strategies, subjects in one quarter of the markets were informed that they could post text messages to each other through the website (one per round in each of the rounds). These 25 markets (subject-pairs) were randomly selected before the experiment started and held fixed throughout. Messages were screened to ensure identities were not revealed in any way.

3 Theoretical predictions

In infinitely repeated games, a wide range of different equilibrium strategy profiles are possible. Since I only occasionally recover one-period-ahead strategies (which I refer to as "strategies" for expositional convenience), it is not feasible to use my data to test whether subjects are actually using *equilibrium* strategies or not. It is always possible that observed one-period-ahead strategies that do not look like they are part of any standard equilibrium are still part of some complicated equilibrium strategy profile. Rather, my focus is on whether elicited strategies have properties that are consistent with the trigger strategies that economists assume in repeated game settings, in which the punishment is immediate, harsh and independent of the nature of the deviation. This covers the Grim and Optimal (symmetric two-phase) strategies described in the Introduction.

To characterize strategies used by subjects to sustain cooperative outcomes, I first define what I mean by cooperative outcomes. The prices (p_1^c, p_2^c) of matched subjects 1 and 2 are said to be cooperative if $\pi_1(p_1^c, p_2^c) \geq \pi^n$ and $\pi_2(p_1^c, p_2^c) \geq \pi^n$, with one of the inequalities strict. Given the payoffs in the experiment, this requires $p_i^c > p^n$ and $\pi_i(p_i, p_j) > \pi^n$ for $i = 1, 2$ and $j \neq i$. This also implies that if $p_1^c \neq p_2^c$, then $|p_1^c - p_2^c| = 1$, $\min(p_1^c, p_2^c) \geq 4$ and $\max(p_1^c, p_2^c) \leq 20$. As will be noted in the next section, subjects in the experiment never set constant but unequal cooperative prices (i.e. $p_{it-1} = p_{it} = p_i^c$ and $p_{jt-1} = p_{jt} = p_j^c$ imply $p_i^c = p_j^c$). This is not surprising given the symmetric design of the experiment. For this reason, in what follows, the focus will be on cooperation in which subjects set symmetric prices.

Denote subject i's one-period-ahead strategy at time t as the function s_{it}, where $s_{it}(p)$ is a mapping from each possible price p that the other subject j could set in round $t-1$ to i's corresponding price in round t. The subscript t indicates that the one-period-ahead strategy that i enters can vary at each point in time, and therefore in practice can be conditioned on the entire history of both i and j's past prices.

The experimental literature on repeated prisoner's dilemma games (e.g. see Fudenberg *et al.*, 2012) has suggested several types of strategies that are distinct in their response to the other subject defecting: *Lenient* (i.e. always cooperate), *Tit-for-Tat* (i.e. match the other subject's previous price), *Grim* (or *Optimal*, as described in the Introduction), and *Always Defect* (i.e. the Non-cooperative strategy). I generalize these here to take into account the wider action set I am considering. I also add the myopic best-response strategy, which would be the same as Always Defect in the prisoner's dilemma game, but differs in the wider action set I am considering. I first summarize the implications for an individual subject's one-period-ahead strategy of each type of strategy.

- *Disproportionate* punishment strategy s_{it}^{D}. If the common cooperative price from i's perspective is $p_i^c > 2$ and the punishment price is $p^p \in \{1,2\}$ then either (i) $s_{it}^{D}(p) = p_i^c$ if $p = p_i^c$ and $s_{it}^{D}(p) = p^p$ if $p \neq p_i^c$, or (ii) $s_{it}^{D}(p) = p_i^c$ if $p \geq p_i^c$ and $s_{it}^{D}(p) = p^p$ if $p < p_i^c$. These are implied by Grim and Optimal in the case that subject i expects subject j to adopt the same cooperative price as himself when deciding which prices to punish, with the subject ignoring upward price deviations in case of (ii). Note $p^p = 2$ for Grim, while $p_i^c = p^m$ and $p^p = 1$ for Optimal.

- *Tit-for-tat* strategy s_{it}^{M}. This involves subject i setting a price equal to subject j's previous price, implying $s_{it}^{M}(p) = p$, $\forall p$.

- *Best-response* strategy s_{it}^{B}. This involves subject i setting a price equal to the one-period (i.e. myopic) best-response $BR(p)$ to subject j's previous price p. Thus, $s_{it}^{B}(p) = BR(p)$, where $BR(p) = p-1$ if $p < 5$ and $BR(p) = p-2$ if $p \geq 5$.

- *Non-cooperative* strategy s_{it}^{N}. This involves subject i setting a price at or lower than the one-shot Nash equilibrium price regardless of subject j's previous price, so $s_{it}^{N}(p) \in \{1,2\}$, $\forall p$.

- *Lenient* strategy s_{it}^{C}. If the common cooperative price from i's perspective is $p_i^c > 2$, this involves $s_{it}^{C}(p) = p_i^c$, $\forall p$. It is consistent with subject i sticking to a cooperative price for (at least) one round regardless of subject j's previous price.

It is straightforward to check that any (common) cooperative price p^c can be supported in equilibrium by both subjects adopting Grim, given the payoff table subjects face and given the 0.95 continuation rate. This also implies that both subjects adopting Optimal (with a sufficiently long punishment phase) also characterizes an equilibrium. Obviously, adopting Non-cooperative with $s_{it}^{N}(p) = 2$, $\forall p$, for $i = 1, 2$ in every round characterizes an equilibrium. This could also be interpreted as part of a punishment phase in case earlier cooperation broke down.

None of Lenient, Best-response or Tit-for-tat as specified can characterize equilibrium strategy profiles in which both subjects adopt these in every round starting from some common cooperative price. This is obvious for Lenient, since then a subject would have an incentive to set the myopic best-response price in every round given the other player sticks to Lenient. For Best-response, if the initial cooperative price is above L$3, then a subject can do better undercutting by L$1 more than would be implied by Best-response (anticipating the other player will also play Best-response). If the initial cooperative price is L$3, then a subject can do better setting L$23 for one round, and then following Best-response in subsequent rounds. For Tit-for-tat, if the initial cooperative price is above L$15, then a subject can do better by lowering her price by L$1 given the price cut will be matched thereafter. If the initial cooperative price is instead between L$3 and L$15, then a subject can do better by setting the price of L$23 for one round, and keeping her price at L$22 thereafter.

4 Experimental results

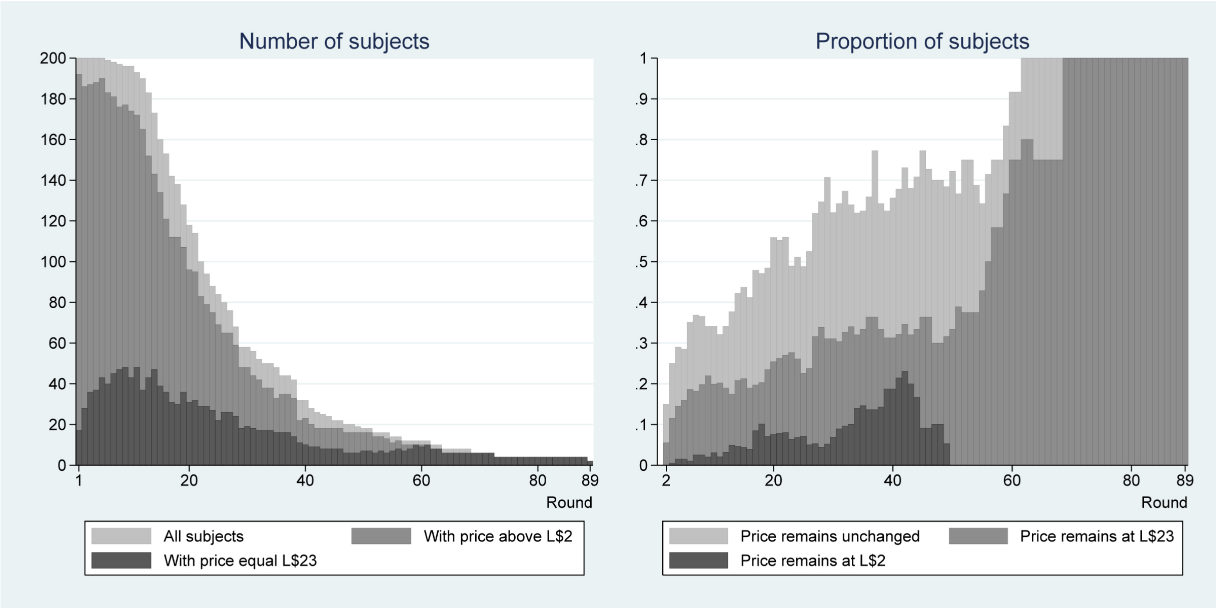
In this section, I present the main results from the experiment. Additional results and alternative specifications are included in the Supplementary Appendix. I first describe the data and some tests of the validity of the experimental design.

4.1 Descriptive statistics

The 200 subjects generated a total of 5,362 price observations (subject – rounds). 204 of these rounds are strategy rounds in which (one-period-ahead) strategies were elicited.[6] Excluding 12 markets for which subjects voluntarily dropped out of the experiment, the remaining 176 subjects generated 4,950 observations. This corresponds to markets lasting 28.1 rounds on average. The average price across all 5,362 observations is L$12.83. The average earnings are S$3.30 per round. This compares favorably with the S$8.74 hourly wage for undergraduate research assistants at the time. Including the S$10 for attending

[6] These came from 132 subjects. Of these, 78 subjects had one strategy elicited, 42 had two strategies elicited, 6 had three strategies elicited and 6 had four strategies elicited.

Figure 1: Number of subjects and proportion holding their prices constant by round



the lab briefing, earnings varied from S$16.10 to S$483.82 excluding the 12 subjects who chose to drop out and obtained only the initial S$10.

The left-hand panel of Figure 1 plots the number of subjects that were active in each round (for all active subjects, those with prices above the one-shot Nash price L$2, and those with prices equal to the monopoly price L$23). As can be seen, some markets last much longer than others. The figure also shows there is an increasing proportion of subjects setting prices above L$2 and in later rounds, equal to L$23. To avoid any bias arising from long-lived subjects contributing more observations to the overall sample, I will focus throughout on statistics that weight each subject equally.

The right-hand panel of Figure 1 shows the proportion of subjects each round that do not change their price from the previous round (for all active subjects, those with a price of L$23 in the previous round, and those with a price of L$2 in the previous round). The proportion of subjects keeping their price constant rapidly increases, so it is about one half of all subjects after 20 rounds. The proportion keeping their price constant at L$23 always exceeds those keeping their price constant at L$2, and beyond 50 rounds no subjects keep their prices constant at L$2.

Subjects in the experiment never set constant but unequal cooperative prices. Once both subjects in a market set the monopoly price, they tend to stick with it (the average price conditional on both subjects setting the monopoly price in the previous round is L$22.95 with only 4.0% of observations involving any change in either of the subjects' prices). In all ten markets in which cooperative prices are set within the first ten rounds and remain constant thereafter, the price was set at the monopoly price.

4.2 Tests of the experimental design

In the Supplementary Appendix I test for various aspects of the experimental design. I find no evidence of a strategy-round effect. Individual subjects set the same prices in their strategy rounds as in their other rounds. Likewise, subjects who are asked for their strategies set the same prices as other subjects who are not asked for their strategies (in the same round).[7] I also find no evidence of a last-round effect in which subjects price differently in their last round. This is not surprising given the randomized stopping rule for each market. One potential weakness of my Internet-based design rather than a more standard lab setting is that even though there is minimal risk that matched subjects communicate (given matched subjects are required to be from different majors), it is difficult to avoid the possibility some non-matched subjects from the same major communicate about the experiment. Based on tests of whether prices or strategies of subjects for a particular major are the same as for all other subjects, I do not find empirical support for a "major effect". Thus, even if some subjects in the experiment (from different markets) do happen to attend the same class and communicate, this does not seem to result in them following a common strategy or setting a common price.

4.3 Do subjects use disproportionate punishment strategies?

In this section, I explore whether subjects support cooperative outcomes through the adoption of strategies with the three main features of disproportionate punishment strategies, namely that they are: (i) immediate, (ii) harsh (e.g. maximal), and (iii) independent of the nature of the deviation. Since the experiment only reveals one-period-ahead strategies, it is only possible to measure the harshness and independence of punishments with respect to the immediate reaction to any deviation. Since Disproportionate involves an immediate punishment following any deviation, this focus on the immediate response to any deviation is still valid for testing whether these strategies are adopted.

A key difference across the different strategies identified in Section 3 is in the harshness of their immediate reaction to any deviation from cooperation. To measure the harshness of any punishment in elicited strategies, I determine the extent of immediate punishment in relation to the maximum possible immediate punishment. Suppose subject i requires that subject j set the price p_j^c in round $t-1$ in order for i to be willing to cooperate and set her highest intended price p_i^c in round t. Then if $p_{jt-1} = p_j^c$, from i's perspective, j has not deviated. On the other hand, different prices (and particularly, lower prices) on the part of j, can be viewed as deviations from the perspective of i. Facing such prices,

[7] The similarity in outcomes between the direct approach and strategy approach is consistent with the findings from most other experiments that compare the two approaches (Brandts and Charness, 2011).

i may set lower prices in response, depending on her punishment strategy. Thus, the harshness of punishment of i's strategy in round t can be measured by:

$$H_{it} = \begin{pmatrix} \frac{\sum_{p=1}^{\overline{p}_{it}-1}(\overline{s}_{it}-s_{it}(p))}{\sum_{p=1}^{\overline{p}_{it}-1}(\overline{s}_{it}-1)} & \text{if } \overline{p}_{it} > 1 \\ 0 & \text{if } \overline{p}_{it} = 1 \end{pmatrix},$$

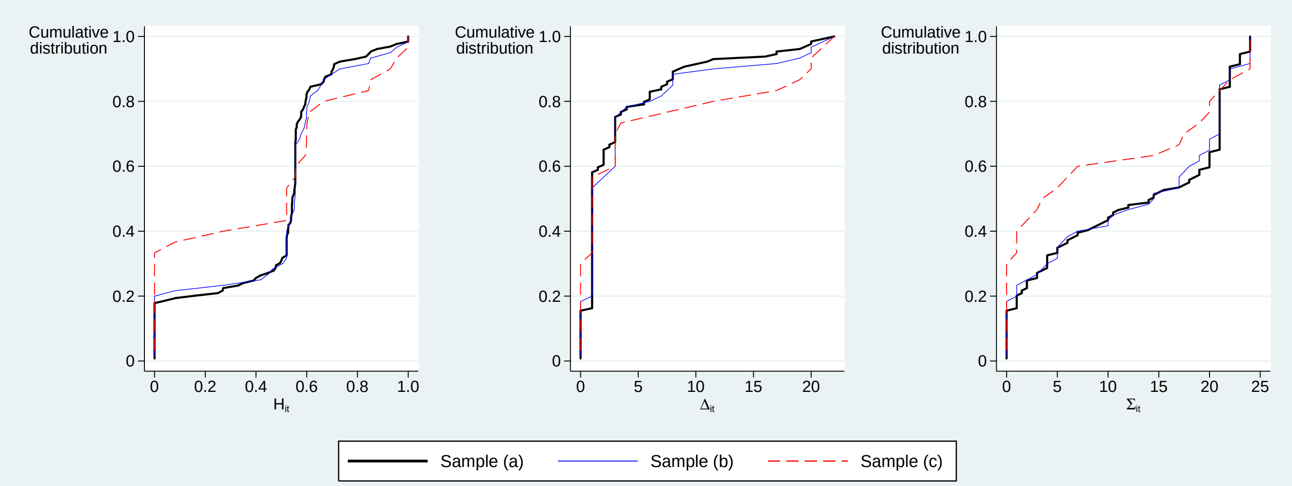
where $\overline{s}_{it} = \max_p s_{it}(p)$ is the maximum price i intends to set in round t and $\overline{p}_{it} = \min[\arg\max_p s_{it}(p)]$ is the lowest price j could set in round $t-1$ that would lead i to set this maximum price in round t.

This harshness index H_{it} captures the extent of a subject's intended punishment as a proportion of the maximum punishment possible (i.e. setting her price at L\$1 for any deviation price). The identifying assumption is that i's maximum intended price $\overline{s}_{it}$ is taken as the cooperative price that i intends to set. Recall $\overline{p}_{it}$ is the lowest price of j in the previous round that still leads i to set the cooperative price $\overline{s}_{it}$. Any price below $\overline{p}_{it}$ is therefore taken as a deviation from the perspective of i and H_{it} captures i's response to these deviations as a proportion of the maximum punishment possible. The index varies between 0 and 1. For example, when both subjects intend to set some cooperative price $p^c > 2$ under Optimal, $H_{it} = 1$ since $\overline{s}_{it} = \overline{p}_{it} = p^c$ and $s_{it}(p) = 1$ for $p < \overline{p}_{it}$. Under Grim instead, $\overline{s}_{it} = \overline{p}_{it} = p^c$ and $H_{it} = (p^c - 2)/(p^c - 1)$, which is close to 1 provided p^c is sufficiently high (e.g. it equals 0.955 for the monopoly price $p^c = 23$). For Tit-for-tat, in which $s_{it}(p) = p$, $H_{it} \approx 0.521$ reflecting a moderate amount of punishment. For Best-response, $H_{it} \approx 0.555$. On the other hand, if i adopts Lenient, then $\overline{p}_{it} = 1$ and $H_{it} = 0$, reflecting that there is no immediate punishment.

The left-hand panel of Figure 2 plots the cumulative distribution of H_{it} for three different subsamples: (a) using the median value of H_{it} for each of the 129 subjects that have at least one elicited strategy involving a maximum intended price above the one-shot Nash equilibrium price, (b) using the same measure but only for the subset of strategies in which both subjects in the market had also set cooperative prices in the previous round (this reduced the sample to 60 subjects), and (c) using the same measure but only for the subset of strategies in which both subjects in the market had set monopoly prices in the previous round (this reduced the sample to 30 subjects). Corresponding to these three different samples, the median H_{it} reported across subjects is 0.543 in case (a), 0.554 in case (b) and 0.521 in case (c).

The cumulative distributions in Figure 2 reveal a bimodal (or possibly trimodal) distribution. Between one fifth and one third of subjects are completely lenient in their immediate response to a deviation by the other subject, depending on the extent of previous cooperation (i.e. depending on which sample (a)-(c) is used).[8] The remaining subjects either use strategies that are similar in their immediate harshness to Tit-for-tat, or are harsher. However, in all samples, only a small percentage use strategies as harsh as implied by Grim. Out of the 129 subjects in sample (a), 96.1% of them have punishment strategies that are less harsh (in their immediate response to a deviation) than implied by Grim for the same maximum intended price. Similarly, for samples (b)-(c), 95% and 90% of strategies are less harsh (in their immediate response) than implied by Grim for the same maximum intended price.

Figure 2: Cumulative distributions of H_{it}, Δ_{it} and Σ_{it} for three different subsamples



Notes: Median values of H_{it} (the harshness index), Δ_{it} (the height of the biggest step of the strategy) and Σ_{it} (the number of upward steps of the strategy) are constructed for each subject from the following subsamples:
(a) 129 subjects that had strategies elicited in which both subjects were willing to set price above p^n
(b) 60 subjects that also had strategies elicited in which both subjects set cooperative prices in the previous round
(c) 30 subjects that also had strategies elicited in which both subjects set monopoly prices in the previous round

To test formally whether the elicited strategies are more or less harsh than implied by Grim, I calculate the difference between H_{it} and its predicted value under Grim. I take the median value of this by subject for the corresponding samples (a)-(c). Using a two-tailed Wilcoxon signed-rank test, I can reject the null hypothesis that this differenced variable is zero at the 1% significance level (p-value is zero to four decimal places) in all three samples (a)-(c). This is not surprising given the median difference between H_{it} and the predicted H_{it} under Grim across the 129 subjects is -0.400, so the typical subject

[8]This compares to the 42% of the eligible histories showing leniency as reported in Fudenberg *et al.* (2011) in their treatment without noise. Leniency in their prisoner's dilemma game requires a subject cooperates despite the other subject defecting in the previous round and both subjects having cooperated in all earlier rounds. Compared to a prisoner's dilemma setting in which subjects face only two extreme choices — cooperate or defect — a single period of leniency in my setting is a very strong requirement. It requires no punishment for the full range of possible deviations and given the full range of possible pricing responses (i.e. punishments). The fact that Fudenberg *et al.* find a higher leniency rate is therefore not surprising, even after allowing for the fact their continuation rate is 0.875, so lower than the 0.95 rate used in this study.

has a one-period-ahead strategy that is much less harsh than implied by Grim.

Another way to measure the extent to which strategies satisfy the properties of disproportionate punishment strategies is to consider s_{it} as a series of steps corresponding to the change in intended price for each increase in p of L$1. Using this approach, an alternative measure of whether punishment is disproportional is the height of the greatest upward step of the strategy. A high step indicates a large decrease in i's intended price for a L$1 decrease in subject j's price (i.e. Disproportionate). Define the height of a step at the price p as $\Delta s_{it}(p) = s_{it}(p+1) - s_{it}(p)$. The height of the greatest upward step of $s_{it}(p)$ is defined as:

$$\Delta_{it} = \max_{1 \leq p \leq 24} \Delta s_{it}(p) .$$

According to Disproportionate, Δ_{it} should be large (equal to $p_i^c - 2$ for Grim and $p_i^c - 1$ for Optimal) but close to 1 for strategies with proportionate or more graduated punishments.

The middle panel of Figure 2 plots the cumulative distribution of Δ_{it} for the same three subsamples (a)-(c). As before, the median value of Δ_{it} by subject is used, calculated for those strategies defined by the respective subsamples. The cumulative distributions reveal that for most subjects, their maximum step-up in intended price is very small (either nothing or L$1), indicating no or only very minimal punishments to a L$1 price reduction of the other subject. Above that, subjects are fairly evenly spread out in their values of Δ_{it}, with only 8.5% of the 129 subjects having $\Delta_{it} > 10$.

To test formally whether the elicited strategies involve different values of Δ_{it} than implied by Grim, I calculate the difference between Δ_{it} and its predicted value under Grim. I take the median value of this by subject for the corresponding samples (a)-(c). Using a two-tailed Wilcoxon signed-rank test, I can reject the null hypothesis that this differenced variable is zero at the 1% significance level (p-value is zero to four decimal places) in all three samples (a)-(c). The median difference between Δ_{it} and the predicted Δ_{it} under Grim is -L$19 for subjects in sample (a) and -L$20 for subjects in samples (b) and (c), reflecting that the typical subject has a one-period-ahead strategy that is proportional with no steps larger than L$1 for L$1 changes in the other subject's price.

Similar to the measure Δ_{it}, according to the independence property of disproportionate punishment strategies, s_{it} should consist of only one (large) step up (i.e. from the punishment price to the cooperative price). In contrast, if the punishment depends on the extent of the crime, then there will be many such small steps. Let $\sigma_{it}(p) = 1$ if $\Delta s_{it}(p) > 0$, at the price p. The number of upward steps of the strategy of subject i in round t is defined as:

$$\Sigma_{it} = \Sigma_{p=1}^{24} \sigma_{it}(p) ,$$

which equals one for Disproportionate but is large for strategies with punishments that depend on the size of the deviation such as Tit-for-tat. For those strategies involving some punishment, Σ_{it} provides a measure of how independent the punishment is.

The right-hand panel of Figure 2 plots the cumulative distribution of Σ_{it} for the same three different subsamples (a)-(c). The median value of Σ_{it} by subject is again used, calculated for those strategies defined by the respective subsamples. The cumulative distributions reveal that most subjects either have no upward steps in their intended price (15.5% out of the 129 subjects in sample (a) had $\Sigma_{it} = 0$) indicating no immediate punishment at all, or many upward steps, with 41.1% of the 129 subjects in sample (a) having 20 or more upward steps and 57.4% of subjects having 10 or more upward steps. In comparison, only 4.7% of subjects in sample (a), 5.0% of subjects in sample (b) and 10% of subjects in sample (c) had only one upward step as implied by Disproportionate.

Figure 2 suggests the properties of elicited strategies are more consistent with Tit-for-tat than Disproportionate, but there remain some subjects (about 5-10%) adopting the latter strategies, as well as a more significant group adopting Lenient.

4.4 Characterizing adopted strategies

In this section, I attempt to characterize directly the (one-period-ahead) strategies elicited according to the different strategies proposed in Section 3 (i.e. Disproportionate, Tit-for-tat, Best-response, Non-cooperative, or Lenient).

Each of the 204 elicited strategies are assigned to one of these five types of strategies, provided the fit is sufficiently close. The closeness of fit between the elicited strategy of subject i in round t and the theoretical strategy k is measured by the root mean square error between the two:

$$RMSE_{it}^{k} = \sqrt{\frac{\sum_{p=1}^{25} \left(s_{it}(p) - s_{it}^{k}(p)\right)^{2}}{25}}. \tag{1}$$

The theoretical strategy k that has the lowest value of $RMSE_{it}^{k}$ is selected provided $RMSE_{it}^{k} \leq \tau$, where τ is a threshold value to ensure that strategies are only assigned if they are a close fit to one of the five types. I consider both the case $\tau = 0$, which requires an exact fit, and the case $\tau = 1$, which requires that a strategy be no more than L\$1 away on "average" (i.e. across p) from the closest theoretical strategy (i.e. a close fit).

Table 1 presents properties on the first elicited strategy of the 132 subjects that had strategies elicited. Best-response, Lenient and Tit-for-tat account for more than 85% of the assigned strategies. These strategies have properties that are in stark contract to Disproportionate — subjects may not respond immediately to a rival's price change (Lenient), or where they do react, the intended price change is roughly proportional to the rival's price change (Tit-for-tat and Best-response). Only two out of 132 subjects had their first elicited strategy classified as Disproportionate when $\tau = 0$ (this increased to seven when $\tau = 1$). In other words, consistent with the earlier evidence looking at the specific features of punishment strategies, the vast majority of subjects do not adopt disproportionate punishment strategies. Evidence from the last (instead of first) elicited strategies and from the path of prices for unassigned strategies (shown in the Supplementary Appendix) supports the same conclusion.

Table 1: Assignment of first elicited strategies and properties for each assignment

		Number assigned	RMSE	Harshness index H_{it}	Greatest upward step Δ_{it}	Number of upward steps Σ_{it}
Based on exact fit	$(\tau = 0)$					
Disproportionate	$p^p = 1$	1	0	1	22	1
	$p^p = 2$	1	0	0.95	21	1
Tit-for-tat		6	0	0.52	1	24
Best-response		19	0	0.55	1	21
Non-cooperative		5	0	-	-	-
Lenient		18	0	0	0	0
Unassigned		82	1.86	0.52	3.90	13.18
Based on close fit	$(\tau = 1)$					
Disproportionate	$p^p = 1$	2	0.22	1	20.50	1
	$p^p = 2$	5	0.56	0.69	8.80	3.20
Tit-for-tat		7	0.14	0.52	1.29	23
Best-response		49	0.31	0.56	1.41	20.92
Non-cooperative		5	0	-	-	-
Lenient		19	0.02	0	0.05	0.11
Unassigned		45	2.95	0.48	4.98	9.33

Notes: Based on the first elicited strategy for each of the 132 subjects that had strategies elicited. Aside from column 1, all other entries are averages of the respective measures across subjects that have first elicited strategies of the corresponding type. The RMSE for unassigned strategies is calculated based on the closest fit strategy in each case.

The finding that very few subjects adopt Disproportionate does not seem to be explained by subjects lacking sufficient experience. In probit regressions of a dummy variable for a subject's adoption of Disproportionate on a constant and the corresponding number of rounds at which the subject was asked for their strategy, the coefficient on the number of rounds is negative and insignificant in all specifications. In contrast, the adoption of Tit-for-tat is positively related to the number of rounds at which a subject is asked for their strategy with the coefficient significant at the 5% level. The estimated probability of adopting Tit-for-tat increases by between 1.5% and 4.0% per round depending on the specification (see the Supplementary Appendix for details).

4.5 Which strategies do better?

The focus thus far has been on characterizing the different (one-period-ahead) strategies chosen by subjects. Having shown the vast majority of subjects do not adopt strategies that are as disproportionate as Grim or Optimal in their immediate response to deviations, it is natural to ask whether this choice comes at a cost to the subjects concerned.

Table 2 shows estimates from regressions of the present discounted value (PDV) of earnings on subjects' choice of strategy. The PDV of earnings is constructed by calculating the present discounted value of all earnings from the current round (in which the strategy is elicited) till the last round for a subject, where discounting is based on the likelihood the market ends each round (0% up to round 12 and 5% thereafter). For rounds beyond the last actual round, the subject's earnings are fixed at their level in their last actual round and the PDV of subsequent earnings (out to infinity) is calculated using the 5% likelihood of markets ending each round. The PDV of earnings takes into account that future earnings are worth less since there is a chance the market will end before they can be realized, and therefore enables me to compare the earnings impact of strategy choices across markets that end in different rounds. Whenever a subject's strategy classification remains the same over multiple strategy rounds, only the first of these observations is used in the regression.

The first two columns of Table 2 differ depending on whether strategies are assigned only if they fit their theoretical counterpart precisely ($\tau = 0$) or with a close fit ($\tau = 1$). Based on either approach, Tit-for-tat is associated with the highest PDV of earnings and Non-cooperative with the lowest PDV of earnings. Given the lack of a constant in the regression, the significance of the coefficients in the first two columns in the top half of Table 2 is uninteresting. What matters is their relative significance. To this end, Wald tests are carried out to test the hypothesis that a strategy has the same effect on the PDV of earnings as another strategy, with p-values reported for each comparison in the bottom half of Table 2. The results imply subjects adopting strategies different from Disproportionate do not obtain significantly different earnings, with the exception that Non-cooperative does worse. In contrast, subjects adopting Tit-for-tat obtain significantly higher earnings than each of the other types of strategies.

One reason a strategy could do better than others is not because it leads to higher earnings, but that higher past earnings explain both the choice of strategy and higher

Table 2: Estimated effect of strategy choice on PDV of earnings

	Model 1		Model 2	
	$\tau = 0$	$\tau = 1$	$\tau = 0$	$\tau = 1$
Disproportionate	84.844***	65.123***	19.641	12.715
	(22.492)	(13.952)	(17.079)	(12.813)
Tit-for-tat	113.832***	89.997***	54.303***	32.627***
	(13.947)	(8.845)	(19.955)	(9.285)
Best-response	47.295***	68.057***	5.824	16.153*
	(7.708)	(6.863)	(10.125)	(8.645)
Non-cooperative	28.557***	34.888***	3.603	7.642
	(8.720)	(9.573)	(5.474)	(7.567)
Lenient	70.700***	62.568***	14.703	3.293
	(10.588)	(10.437)	(12.039)	(11.756)
Unassigned	66.136***	58.888***	16.523*	5.600
	(5.783)	(7.419)	(8.484)	(9.771)
past earnings per round			15.539***	16.751***
			(2.586)	(2.451)
p-values (Wald tests of equal coefficients)				
Ho: Disproportionate = Tit-for-tat	0.275	0.135	0.083*	0.049**
Ho: Disproportionate = Best-response	0.117	0.849	0.353	0.755
Ho: Disproportionate = Non-cooperative	0.021**	0.076*	0.287	0.681
Ho: Disproportionate = Lenient	0.570	0.884	0.735	0.463
Ho: Disproportionate = Unassigned	0.450	0.682	0.820	0.521
Ho: Tit-for-tat = Best-response	0.000***	0.040**	0.006***	0.017**
Ho: Tit-for-tat = Non-cooperative	0.000***	0.000***	0.006***	0.004***
Ho: Tit-for-tat = Lenient	0.015**	0.047**	0.026**	0.002***
Ho: Tit-for-tat = Unassigned	0.002***	0.008***	0.026**	0.001***
Ho: Best-response = Non-cooperative	0.110	0.006***	0.798	0.342
Ho: Best-response = Lenient	0.077*	0.663	0.401	0.227
Ho: Best-response = Unassigned	0.044**	0.347	0.197	0.227
Ho: Non-cooperative = Lenient	0.003***	0.052*	0.294	0.710
Ho: Non-cooperative = Unassigned	0.001***	0.052*	0.065*	0.838
Ho: Lenient = Unassigned	0.700	0.770	0.843	0.833

Notes: Dependent variable is PDV of earnings as measured in Singapore dollars (S\$1≈US\$0.70). Whenever a subject's strategy classification remains the same over multiple strategy rounds, only the first of these observations is used in the regression. There are 151 such observations when $\tau = 0$ and 153 observations when $\tau = 1$. Robust standard errors, clustered on subjects, are reported in parentheses and used in the Wald tests.

*** Significant at 1%. ** Significant at 5%. * Significant at 10%.

future earnings. For instance, Non-cooperative is more likely to arise if earnings have been low in the past, and such earnings are likely to persist. If this is true, then strategy choices should no longer affect the PDV of earnings once past earnings are controlled for. Columns 3 and 4 repeat the estimation and tests in columns 1 and 2 once a control for past earning is added (i.e. the average of a subject's past earnings per round).

The estimates in columns 3 and 4 show that even controlling for past earnings, the PDV of earnings is significantly higher when subjects adopt Tit-for-tat than any other strategies including Disproportionate. Non-cooperative no longer does significantly worse in terms of earnings than the remaining strategies, with the exception of Unassigned (but only at the 10% significant level and only for $\tau = 0$). Other differences in earnings between strategies are no longer significant. The results suggest Tit-for-tat leads to higher earnings (the increase ranges from S\$19.91 or about US\$14 more than Disproportionate with $\tau = 1$ to S\$34.66 or about US\$24 more than Disproportionate with $\tau = 0$).

That Tit-for-tat does best in this analysis is reminiscent of the findings of Selten *et al.* (1997), as well as the earlier literature (e.g. Axelrod and Hamilton, 1981) on the success of Tit-for-tat in tournaments involving repeated prisoner's dilemma games. Given that paired subjects only adopt identical one-period-ahead strategies in 6.86% of cases, a tournament approach may be relevant for understanding this finding. This is particularly true if subjects find it difficult to work out each other's strategies, which is likely to be the case when experimentation is very costly. This may also explain why Lenient does not lead to significantly lower earnings compared with Best-response, Disproportionate or Unassigned strategies. For subjects that have already coordinated on monopoly prices, the purpose of doing so may be to avoid reacting to an accidental price cut of the rival.

For example, in one case a subject that used Lenient confirmed in his message that he put the price of L\$23 in round 10 (a strategy round) since he assumed any other price of the other subject "was probably a mistake". In another case, the subject writes "Hey I put 22 for this round cos you put 22 for the last? I'll assume it's a typo cos you could have got more if u put 21", and the other subject from the same market writes "sorry, last round I was too excited about the number 22, so accidently entered 22 for the price". These two subjects maintained a price of L\$23 in all rounds from 2 to 46 except for the round in which one set the price of L\$22 and the subsequent round when the other set a price of L\$22 in response.

4.6 The role of communication

A detailed analysis of how prices and strategy choice depend on whether communication is allowed is contained in the Supplementary Appendix. Here I just summarize the main findings. As with most other studies, I find prices are significantly higher for subjects in the group where communication is allowed (this is true for most of the rounds, based on a Wilcoxon rank-sum test between the two groups). A natural question to ask is whether this is because communication allows subjects to adopt certain types of strategies?

The main role of communication seems to be to facilitate coordination on a high price. There is very little communication about subjects' intentions to punish deviations in the 239 messages sent between subjects that are allowed to communicate. The majority of messages are either proposals to set higher prices, affirmations of the intention to stick to a high price, positive comments noting how much money subjects were both earning, or general comments noting that their payoffs will be high (low) if they both set high (low) prices. Indeed, 110 of the 239 messages contain the monopoly price "23". Threats of punishment (low prices) in case high prices are not adopted are mentioned in only six of the 239 messages (corresponding to five subjects).

Probits are conducted to see if the adoption of a particular strategy is more likely under the communication treatment. I find subjects are significantly more likely to adopt Lenient and less likely to adopt Non-cooperative when allowed to communicate. There are no significant results for other strategies. Focusing instead on subjects that had actually communicated prior to strategy elicitation, there is some weak evidence that subjects are also more likely to adopt Disproportionate in this case and less likely to adopt Best-response (although it is not robust across specifications).

Our results are consistent with communication being used to support cooperative strategies rather than "undercutting" or non-cooperative strategies. The evidence from the subjects' messages suggests that this mainly arises through subjects communicating the benefits of cooperation rather than discussing specific punishment strategies directly. Reflecting this, the main findings of the paper are robust to whether the communication treatment is included or not.

5 Conclusion

In this paper I designed an experiment to see the extent to which subjects adopt certain types of strategies in a repeated competition setting. Rather than trying to make inferences from prices alone, I provided direct evidence by also occasionally eliciting subjects' one-period-ahead strategies. I found the elicited strategies often involve no immediate punishment following a deviation from cooperation (they are lenient). Where punishment is immediate, it is not as harsh as the Grim strategy, and indeed the severity of the punishment depends on the extent of the other subject's deviation.

These properties cast doubt on the use of disproportionate punishment strategies assumed in much of the applied theory literature (e.g. the literature on collusion). In the few cases where subjects adopted such strategies, their rivals adopted different types of strategies. This is despite the fact that subjects faced a symmetric environment in which the monopoly price was the obvious cooperative outcome to coordinate on, and in which shocks were absent. In more realistic environments, such as those faced by competing firms in real world markets, coordinating on these types of equilibria would seem to be much more challenging. A possible exception is if decision makers can communicate extensively with each other, as could be the case in overt collusion cases involving cartels. However, I do not find disproportionate punishment strategies are used significantly more often in my communication treatment, possibly reflecting the fact that I only allow for limited communications (one message per round).

The evidence provided in this paper suggests Tit-for-tat is the most profitable strategy for subjects, even controlling for their past payoffs. This is broadly consistent with evidence from the large literature on repeated prisoner's dilemma game tournaments (starting with Axelrod and Hamilton, 1981) showing Tit-for-tat performs well. In a sense, my findings generalize these previous results since in the present game Tit-for-tat involves matching any previous price of the other subject (so the action space is much richer) and the setting is a standard infinitely repeated game setting.

More generally, the results suggest softer punishment strategies than Grim are typically used by subjects, and are at least as profitable. Since I can only observe one-period-ahead strategies, this can be for one of two different reasons — that subjects do not make full use of the available punishments following a deviation, or that subjects tend to not react at all initially, although they may resort to maximal punishment if the deviation persists. The evidence given in this paper suggests both types of punishment are likely to be important.

One possible explanation for these results is that subjects have non-strategic motivations. It could be that some subjects are just more cooperative and trusting than others, and are willing to give others the benefit of the doubt in case of a one-time deviation. I tried to minimize the role of social preferences by framing the experiment as a market game, in which each subject represents a seller that sets a price. As discussed in the introduction, recent evidence casts doubt on the role of social preferences in driving cooperation in this type of repeated competition setting.

An alternative explanation for the results I find is that "soft" punishment strategies avoid an escalation of conflict that may arise from accidental deviations or misunderstandings due to strategic uncertainty, while still providing a reasonable level of punishment to help discipline deviations. When reaching (or returning to) the cooperative outcome is difficult subjects may be better off using a softer punishment initially to avoid getting stuck at very bad outcomes. Additionally, Tit-for-tat may give a more salient signal of punishment than other approaches, since the subject's deviation is precisely matched. Determining what is actually driving subjects to choose "soft" punishments, and building a theory based on it, is an important and challenging future direction for the theory of cooperation in repeated games.

References

Abreu, D. (1986) "Extremal equilibria of oligopolistic supergames," *Journal of Economic Theory* 39: 191-225.

Abreu, D. (1988) "On the theory of infinitely repeated games with discounting," *Econometrica* 56: 383-396.

Axelrod, R. and W. Hamilton (1981) "The evolution of cooperation," *Science* 21: 1390-1396.

Brandts, J. and G. Charness (2011) "The strategy versus the direct-response method: A first survey of experimental comparisons," *Experimental Economics,* 14: 375-398.

Cabral, L., E. Ozbay and A. Schotter (2012) "Intrinsic and instrumental reciprocity: An experimental study," Working Paper.

Camera, G., and M. Casari. 2009. "Cooperation among strangers under the shadow of the future," *American Economic Review*, 99(3), 979-1005.

Camera, G., M. Casari and M. Bigoni (2012) "Cooperative strategies in anonymous economies: An experiment," *Games and Economic Behavior*, 75: 570586.

Dal Bó, P. and G. R. Fréchette (2011) "The evolution of cooperation in infinitely repeated games: Experimental evidence," *American Economic Review*, 101: 411-429.

Dal Bó, P. and G. R. Fréchette (2012) "Strategy choice in the infinitely repeated prisoners dilemma," Working paper.

Dreber, A., D. Fudenberg, and D. G. Rand (2011) "Who cooperates in repeated games?" mimeo. Harvard University.

Engel, C. (2007) "How much collusion? A meta-analysis of oligopoly experiments," *Journal of Competition Law and Economics*, 3(4), 491-549.

Engle-Warnick, J. and R. L. Slonim (2006) "Inferring repeated-game strategies from actions: Evidence from trust game experiments," *Economic Theory*, 54: 95-114.

Friedman, J. (1971) "A noncooperative equilibrium for supergames," *Review of Economic Studies* 38: 1-12.

Friedman, J. and L. Samuelson (1990) "Subgame perfect equilibrium with continuous reaction functions," *Games and Economic Behavior* 2: 304-324.

Friedman, J. and L. Samuelson (1994) "Continuous reaction functions in duopolies," *Games and Economic Behavior* 6: 55-82.

Fudenberg, D., D. G. Rand, and A. Dreber (2012) "Slow to anger and fast to forgive: Cooperation in an uncertain world," *American Economic Review*, 102: 720-749.

Garrod, L. (2012) "Collusive Price Rigidity under Price-Matching Punishments," *International Journal of Industrial Organization*, 30: 471-482.

Kalai, E. and W. Stanford (1985) "Conjectual-variations strategies in accelerated cournot games," *International Journal of Industrial Organization* 3: 133-152.

Lu, Y. and J. Wright (2010) "Tacit collusion with price-matching punishments," *International Journal of Industrial Organization*, 28: 298-306.

Mason, C.F. and O.R.Phillips (2002) "In support of trigger strategies: Experimental evidence from two-person noncooperative games," *Journal of Economics and Management Strategy* 11: 685-716.

Reuben, E. and S. Seutens (2012) "Revisiting strategic versus non-strategic cooperation," *Experimental Economics* 15: 24-43.

Samuelson, L. (1987) "Nontrivial subgame perfect duopoly equilibria can be supported by continuous reaction functions," *Economics Letters* 24, 207-211.

Selten, R. (1967) Die strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eimes Oligopolexperiments. In: Sauerman, H. (Ed.), Beiträge zur experimentellen Wirtschaftsforschung. J.C.B. Mohr (Paul Siebeck), Tübingen, pp. 136–168.

Selten, R., Mitzkewitz, M. and G. R. Uhlich (1997) "Duopoly strategies programmed by experienced players," *Econometrica* 65: 517-555.

Slade, M.E. (1992) "Vancouver's gasoline-price wars: An empirical exercise in uncovering supergame strategies," *Review of Economic Studies* 59: 257-276.