Essays in Behavioral and Experimental Economics

DISSERTATION

ZUR ERLANGUNG DES AKADEMISCHEN GRADES
DOCTOR RERUM POLITICARUM

AN DER

FAKULTÄT FÜR WIRTSCHAFTS- UND SOZIALWISSENSCHAFTEN
DER RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG

VORGELEGT VON

Peter Dürsch

HEIDELBERG, OKTOBER 2010

Table of contents

Introduction

Behavioral and experimental economics are two relatively recent and closely related fields in economics. Experimental economics adds experiments as a method of research to theoretical modeling, empirical analysis of real world data and simulations. This method is not specific to any field of economics, and experiments have been used for a long time, but it is in connection with behavioral economics that experiments have become more refined and often used. Behavioral economics relaxes two important assumptions that are at the core of almost all economic modeling: That humans are rational and selfish (money maximizing). After giving up these, experimental methods are used to answer the question: If not rational and selfish, what else instead?

Relaxing the rationality assumption questions the assertion that humans behave as if they were strong and flawless computers, capable of performing calculations of arbitrary difficulty instantly, without ever making a mistake. And indeed, there are many results that show humans to be only boundedly rational, or to be subject to biases that deviate from rationality. The first part of the dissertation falls into this branch of behavioral economics. Learning theories postulate that, in dynamic decision situations, humans use rules of thumb, based on the observable history, to help them decide. We test some computerized versions of learning theories that try to describe human behavior, and find that, with one exception, they are rather easily manipulated by their (human) opponents.

A second branch of behavioral literature stems from a weakening of the second assumption, selfishness. That is, humans try not only to maximize their own outcome, but also care for the effects their behavior has on other humans. This is commonly called other-regarding preferences, or social preferences. In the second part of the dissertation, we look at a special form of other-regarding preferences, punishment. In a one shot, anonymous game punishment does not serve any monetary purpose. Yet, in experiments, subjects use punishment, even if it is costly and they themselves can not derive any profit from punishing. We investigate punishment when it is risky, and whether subjects have a desire to punish personally.

In the first chapter, we explore how human subjects placed in a repeated Cournot duopoly react to opponents who play according to five popular learning theories: (Myopic) Best Response, Fictitious Play, Reinforcement Learning, Trial & Error and Imitate-the-Best. All of these have been proposed in the literature as theories to model and describe the behavior of humans. The usual test of these models in the literature is to measure real human behavior (for example in laboratory experiments) and then to fit the learning theories to the observed behavior. We turn around the stick and ask: If someone indeed behaved according to these theories, how would others react, and how successful would he be? To achieve this, we program computer algorithms that play according to the above learning theories and let subjects play against these algorithms. The main experiment was implemented as an internet study, enabling us to recruit a diverse set of subjects who played outside of the usual, artificial, laboratory setting. However, we also include a laboratory treatment to check for qualitative differences.

Despite not being informed about the specific learning theory they are matched with, our subjects prove to be surprisingly successful. They achieve high average payoffs and, importantly, higher payoffs than the learning theories. The only exception to this are subjects who are matched with Imitate-the-Best. Looking at the learning theories used, it turns out that all but Imitate-the-Best can be induced to play low, accommodating quantities in later round by aggressive, high-quantity play by the human subjects in early rounds. These early high quantities lead to up-front lower profits but are rewarded by higher long term profits. Imitate-the-Best is the only algorithm that can not be influenced in this way. We conclude that subjects are not merely playing myopically in each round of the repeated game, but "strategically teach" their opponents to play in a way that raises future own profits.

While the first chapter is rather explorative, the second part of the dissertation looks at a topic that has already received considerable attention in the literature: Punishment by peers. We investigate punishment in two special cases, direct, personal punishment, and punishment as a risky instrument.

Chapter two is concerned with the way punishment is enacted. Do punishing subjects seek only a decrease in the well-being of the "offender", or do they want to personally bring that decrease about, do they want to be involved in the act of punishment? Subjects having such a desire to punish personally, instead of punishment being enacted by someone else, would imply that the way punishment is institutionalized, e.g. in justice systems where punishment is enacted by state employees, will have an impact on the utility of those who were wronged.

We implement punishment in a design where the desire to punish personally is separated from other potential incentives. Subjects bid for the right to be the ones to punish in a second price auction. Bidders can neither affect the probability or strength of punishment, which is fixed earlier in the game, nor can they send a signal to the offender. The act of punishment is represented by physical destruction of a part of the offender's allocation. While at first sight the results seem to indicate that subjects are willing to spend money to win the right to punish personally, that view is tempered by a control treatment which consists of an auction alone, without punishment or other monetary prize. In the control subjects do not bid less compared to the main treatment. Therefore, at least for the form of punishment we implement in the lab (of course, the experiment can not include physical harm to the offender), we do not find evidence for a desire to punish personally.

The main question of chapter three is the interaction between risk and punishment. It is well known that many subjects are not risk neutral. At the same time, many subjects show other-regarding preferences, e.g. by engaging in costly punishment. When other-regarding preferences are modeled, risk aversion is typically not taken into account at all, despite the fact that punishment need not always happen under conditions where outcomes are certain. We look at possible interactions in a one-shot prisoner's dilemma game with punishment opportunity. In one treatment, punishment is certain, while in another treatment, the outcome of punishment is subject to a lottery. At the same time, we measure risk aversion via a Holt-Laury test.

Chapter four looks at the similar question of changes in cooperation rates in the prisoner's dilemma for risk-averse subjects, conditional on punishment being present in the design. Both papers are based on the same experimental data and suffer from a lack of instances of punishment happening. To create enough data-points, we tried to maximize both the number of punishment worthy defection-cooperation pairs and subsequent punishment. While we achieved many defection-cooperation pairs, subjects only rarely punish. This might be due to the fact that we use a one-shot prisoner's dilemma or the parameterization of our experiment. For the question of cooperation behavior, this explains the unchanged behavior of subjects we find, if subjects correctly predicted the low amount of punishment. Regarding punishment under risk, the results rely only on a very restricted dataset, but point in the direction that subjects are not impacted by risk on other's payoff as they are by risk in their own payoff.

I. Rage Against the Machines: How Subjects Play Against Learning Algorithms*

Abstract

We use a large-scale internet experiment to explore how subjects learn to play against computers that are programmed to follow one of a number of standard learning algorithms. The learning theories are (unbeknown to subjects) a best response process, fictitious play, imitation, reinforcement learning, and a trial & error process. We explore how subjects' performances depend on their opponents' learning algorithm. Furthermore, we test whether subjects try to influence those algorithms to their advantage in a forward-looking way (strategic teaching). We find that strategic teaching occurs frequently and that all learning algorithms are subject to exploitation with the notable exception of imitation.

*Paper co-authored by Albert Kolb, Jörg Oechssler, Burkhard Schipper

1 Introduction

In recent years, theories of learning in games have been extensively studied in experiments. The focus of those experiments was primarily the question which learning theories describe best the average behavior of subjects. It turns out that some very simply adaptive procedures like reinforcement learning, best response dynamics, or imitation are fairly successful in describing average learning behavior of subjects in some games (see e.g. Erev and Haruvy, 2008, for a recent survey).

The focus of the current experiment is different. We are interested in the following strategic aspects of learning in games. First, how is a player's success affected by the way opponents learn? Second, how can the opponent's learning process be influenced by the player's behavior? For example, can it be manipulated to the player's advantage? To address those questions, we present here a first - very exploratory - experimental study. Since we are interested in how subjects respond to certain learning theories, we need to be able to control the behavior of opponents. The best way to do that is by letting subjects play against computers programmed with particular learning theories.[1]

The questions raised in this paper seem to be fairly novel, although the second question has received some attention at least in the theoretical literature.[2] For example, Fudenberg and Levine (1998, p. 261) write "A player may attempt to manipulate his opponent's learning process and try to 'teach' him how to play the game. This issue has been studied extensively in models of 'reputation effects', which typically assume Nash equilibrium but not in the context of learning theory." Following Camerer and Ho (2001) and Camerer, Ho, and Chong (2002) we shall call this aspect of learning "strategic teaching".[3] We believe that this hitherto largely neglected aspect of learning is of immense importance and deserves further study. As we shall see in this experiment, theories just based on adaptive processes will not do justice to the manipulation attempts of subjects.

We consider five learning theories in a Cournot duopoly: best-response (br), fictitious play (fic), imitate-the-best (imi), reinforcement learning (re), and trial & error (t&e). Some noise is added in order to make the task less obvious. Noise is also a requirement for some

[1] Subjects are, of course, being told that they play against computers. There is now a large experimental literature making use of computers to control for some players' behavior in strategic situation. See Cason and Sharma (2007) for a recent experiment.

[2] See Fudenberg and Levine (1989) and Ellison (1997).

[3] Note, however, that we use the term in a broader sense, not necessarily referring to EWA as in Camerer et al. (2002).

of the theoretical predictions to work as it prevents a learning process from getting stuck at states which are not stochastically stable.[4] The selection of learning theories is based on the prominence in the literature, convenient applicability to the Cournot duopoly, and sufficient variety of theoretical predictions.

The experiment was conducted as a large scale internet experiment. Internet experiments are still relatively novel (see e.g. Drehmann, Oechssler, and Roider, 2005, for first experiences). Arguably, the setting (working at home or in the office at your own PC) is more representative of real world decisions than in the usual laboratory experiments. Also, internet experiments allow to reach a large subject pool at moderate cost.[5]

With respect to the first question, we find that subjects achieve substantially higher profits than all of their computer opponents but one. The exception is the imitation algorithm, for which we show theoretically that it cannot be beaten by more than a small margin and which in fact performs on average better than its human opponents in the experiment. The computer opponent that allows for the highest profits for its human counterparts is the reinforcement learning computer. However, due to the stochastic nature of reinforcement learning, a lot of luck is needed, and the variances are high.

This leads us to the second question: We find that strategic teaching occurs frequently and that all learning algorithms are subject to exploitation with the notable exception of imitation. Subjects learn quickly how to exploit the best response– and trial & error–computers, usually by behaving as Stackelberg leader, although some subjects manage to find more innovative and even more profitable ways.

Two papers are closely related to our work. Shachat and Swarthout (2002) let subjects play against both human subjects and computers, which are programmed to follow reinforcement learning or experienced weighted attraction in repeated 2x2 games with a unique Nash equilibrium in mixed strategies. They find that human play does not significantly vary depending on whether the opponent is a human or a programmed learning algorithm. In contrast, the learning algorithms respond systematically to non-Nash behavior of human subjects. Nevertheless, these adjustments are too small to result in significant payoff gains. Coricelli (2005), on the other hand, found that human subjects do manage to exploit computer opponents that play a biased version of fictitious play in repeated 2x2 zero-sum games.

[4] See e.g. Vega-Redondo (1997) for imitate-the-best and Huck, Normann, and Oechssler (2004a) for trial & error.

[5] Since internet experiments are relatively novel, we explore some methodological issues of this experiment in a companion paper by comparing it to various laboratory treatments (see Duersch et al., 2008).

The remainder of the paper is organized as follows. Section 2 describes the Cournot game that is the basis for all treatments. In Section 3 we introduce the computer types and the associated learning theories. The experimental design is explained in Section 4, followed by the results in Section 5. In Section 6 we consider a laboratory treatment as a robustness check. Section 7 concludes. The instructions for the experiment and screenshots are shown in the Appendix.

2 The Cournot game

We consider a standard symmetric Cournot duopoly with linear inverse demand function $\max\{109 - Q, 0\}$ and constant marginal cost, $MC = 1$. Each player's quantity q_i, $i = 1, 2$ is an element of the discrete set of actions $\{0, 1, ..., 109, 110\}$. Player i's profit function is given by

$$\pi(q_i, q_{-i}) := (\max\{109 - q_i - q_{-i}, 0\} - 1)\, q_i. \tag{1}$$

Table 1 shows outputs and profits for the Nash equilibrium, the competitive outcome (where $p = MC = 1$), the collusive outcome, the Stackelberg outcome, and the monopoly solution. Subjects play the Cournot duopoly repeatedly for 40 rounds. Thus, we index the quantity q_i^t by the period $t = 1, ..., 40$.

Table 1: Prominent outcomes

	q_i	q_{-i}	π_i	π_{-i}
Cournot Nash equilibrium	36	36	1296	1296
symmetric competitive outcome	54	54	0	0
symmetric collusive outcome	27	27	1458	1458
Stackelberg leader outcome	54	27	1458	729
Stackelberg follower outcome	27	54	729	1458
monopoly solution	54	0	2916	0

A Cournot duopoly is chosen for this experiment because, based on earlier theoretical and experimental contributions, we expected that the behavior of the various learning theories would differ in interesting ways in a Cournot game. In particular, there was the conjecture that imitation would behave very differently from the remaining learning theories. In order to make this conjecture precise, we derive in this paper a new theoretical result, namely that the imitation algorithm cannot be beaten by much even by a very sophisticated player. Of course, this result applies only to a particular class of games that includes the Cournot game but also games like chicken.[6]

3 Computer types

Computers were programmed to play according to one of the following decision rules: Best-response (br), fictitious play (fic), imitate the best (imi), reinforcement learning (re), or trial & error (t&e). All decision rules except reinforcement learning are deterministic, which would make it too easy for subjects to guess the algorithm (as we experienced in a pilot study to this project). Therefore, we introduced some amount of noise for the deterministic processes (see below for details). The action space for all computer types was $\{0, 1, ..., 109\}$.

All computer types require an exogenously set choice for the first round as they can only condition on past behavior of subjects. To be able to test whether starting values matter, we chose different starting values. However, to have enough comparable data, we restricted the starting values to 35, 40, and 45. Starting quantities were switched automatically every 50 plays in order to collect approximately the same number of observations for each starting quantity but subjects were unaware of this rule.

3.1 Best-response (br)

Cournot (1838) himself suggested a myopic adjustment process based on the individual best-response

$$q_i^t = \arg\max_{q_i} \pi(q_i, q_{-i}^{t-1}) = \max\left\{\frac{108 - q_{-i}^{t-1}}{2}, 0\right\}, \tag{2}$$

for $t = 2, ...$ Moreover, the parameters are such that if both players use the best-response process, the process converges to the Nash equilibrium in a finite number of steps (see e.g. Monderer and Shapley, 1996).

This deterministic process is supplemented by noise in the following way. If the best response process yields some quantity q_i^t, the computer actually plays a quantity chosen from a Normal distribution with mean q_i^t and standard deviation 2, rounded to the next integer in $\{0, 1, ..., 109\}$.[7] This implementation of noise is also used for computer types fictitious play and imitation.

[6] We thank a referee for this observation.

[7] Due to a programming error in the rounding procedure, the noise of computer types br, fic, and imi was actually slightly biased downwards (by 0.5), which makes the computer player slightly less aggressive. This does not have any lasting effects for computer types br and fic but has an effect on imi.

3.2 Fictitious play (fic)

A second decision rule that is studied extensively in the literature is fictitious play (see Brown, 1951, Robinson, 1951, and Fudenberg and Levine, 1998, chapter 2). A "player" (that is, in our setting, the computer) who uses fictitious play chooses in each round a myopic best response against the historical frequency of his opponent's actions (amended by an initial weight for each action). If we let those initial weights be the same for each action and each player, $w_i^1(q_{-i}) = w^1$, we obtain the following recursive formulation for the weight player i attaches to his opponent's action q_{-i}, where 1 is added each time the opponent chooses q_{-i}.

$$w_i^t(q_{-i}) = w_i^{t-1}(q_{-i}) + \begin{cases} 1 & \text{if} \quad q_{-i}^{t-1} = q_{-i} \\ 0 & \text{if} \quad q_{-i}^{t-1} \neq q_{-i} \end{cases}$$

for $t = 2, ...$ Player i assigns probability

$$p_i^t(q_{-i}) = \frac{w_i^t(q_{-i})}{\sum_{q'_{-i}} w_i^t(q'_{-i})}$$

to player $-i$ using q_{-i} in period t. Consequently, player i chooses a quantity that maximizes his expected payoff given the probability assessment over the opponent's quantities, i.e.,

$$q_i^t \in \arg\max_{q_i} \sum_{q_{-i}} p_i^t(q_{-i})\pi(q_i, q_{-i}). \tag{3}$$

We simulated the fictitious play processes against itself and some other decision rules for many different initial weights w^1 and ended up choosing $w^1 = 1/25$. Except for much smaller or much larger initial weights, results of the simulations did not change much. Very high initial weights lead to rather slow adaptation whereas very small ones resulted in erratic movements. Since our Cournot duopoly is a potential game, fictitious play converges to the unique Cournot Nash equilibrium (see Monderer and Shapley, 1996).

3.3 Imitate the best (imi)

Imitation has received much attention recently in both theory and experiments (see e.g. Vega-Redondo, 1997, Apesteguia et al. 2007, Schipper, 2008). The rule "imitate the best" simply requires to choose the best action that was observed in the previous period. If player i follows this decision rule in $t = 2, ...$, he chooses

$$q_i^t = \begin{cases} q_i^{t-1} & \text{if} \quad \pi(q_i^{t-1}, q_{-i}^{t-1}) \geq \pi(q_{-i}^{t-1}, q_i^{t-1}) \\ q_{-i}^{t-1} & \text{otherwise.} \end{cases} \tag{4}$$

Vega-Redondo (1997) shows for a symmetric Cournot oligopoly that if all players follow this decision rule up to a small amount of noise, then the long run distribution over quantities assigns probability 1 to the competitive outcome (where $p = MC$) as the noise vanishes.[8] The intuition for this is simple. For all total quantities Q such that $p > 1$, the firm with the highest quantity receives the highest profit. When the highest quantity is imitated by other firms, Q increases. For Q such that $p < 1$, the firm with the lowest quantity is being imitated, such that Q decreases. Thus, Q converges to the competitive quantity. See Huck, Normann, and Oechssler (1999) and Offerman, Potters, and Sonnemans (2002) for experimental evidence.

With respect to the current experiment, of particular interest is the question what happens when only one player imitates the best? The following proposition shows that imi can essentially not be exploited even by very sophisticated players.[9]

Proposition 1 *Suppose a player knows that his opponent follows the rule imitate the best (*imi*) and even knows the opponent's starting value.*

(a) If this player wants to maximize his absolute *payoff over 40 rounds, then the optimal strategy yields an average payoff of* 46 374, *which is much less than the profit of* 55068 *for his computer opponent imi and also less than the Cournot profit of* 51 840*.*

(b) If the player wants to maximize his relative *payoff (i.e. the difference between his payoff and his opponent's payoff), then an optimal strategy yields an average profit differential of 212.67.*

Proof: See appendix.

The implications of the proposition are the following. First, even sophisticated human players that maximize their long–run profits in a forward looking way against imi will achieve much lower payoffs than their computer opponent imi. Second, the total profit of the sophisticated player is much lower than the total profit of the stage game Cournot Nash equilibrium over 40 rounds or the profit a Stackelberg leader could achieve against computer br. Hence, even sophisticated human subjects playing against imi will typically

[8] Vega-Redondo's result has been generalized to larger classes of games by Alos-Ferrer and Ania (2005), Possajennikov (2003), Schipper (2003) and Tanaka (1999).

[9] Schipper (2008) shows that if there are both imitators and best–response players in the game, then any state where imitators are weakly better off than best–response players and where best-response players play a best–response is absorbing. Moreover, if mistakes are added, then in the long run imitators are strictly better off than best–response players. The intuition is that if imitators play a sufficiently large quantity, best–responders become Stackelberg followers.

earn much less than playing against other computer types. Third, since imi will never lower its quantity q_2 as long as $q_1 + q_2 < 108$, any mistake or experimentation that leads to an increase of the human subject's quantity will result in a permanent drop of profits. Thus, we should expect to see profits decline over time. Finally, even subjects that do not care to maximize their absolute profits but instead aim at beating the computer (i.e. maximize their relative payoff), can only do so by a very modest margin of 212 (for comparison note that a Stackelberg leader gains a profit differential against a follower of $40 * 729 = 29\,160$).

3.4 Reinforcement learning (re)

In a standard model of reinforcement learning by Roth and Erev (1995), an action is chosen with probability that is proportional to the *propensity* for this action. Propensities, in turn, are simply the accumulated payoffs from taking this action earlier in the process.

In games with a large action space such as a Cournot duopoly, it seems unreasonable to reinforce only that single action that was chosen in a given round. Rather, actions in the neighborhood should also be reinforced although to a lesser extent depending on their distance to the original choice. Therefore, we complement the standard model of reinforcement learning by updating of neighborhoods à la Sarin and Vahid (2004).

The player starts with the same initial propensity for each quantity, $w_i^1(q)$. For $t = 2, \ldots$propensities are updated by[10]

$$w_i^t(q) = \max\left[1, w_i^{t-1}(q) + \beta(q, q_i^{t-1})\pi(q_i^{t-1}, q_{-i}^{t-1})\right],$$

where β is the linear Bartlett function

$$\beta(q, q_i^{t-1}) := \max\left\{0, \frac{6 - |q - q_i^{t-1}|}{6}\right\}.$$

That is, all actions within 5 grid points of the chosen action are also reinforced.

The probability of playing quantity q in period t is computed by normalizing the propensities

$$p_i^t(q) = \frac{w_i^t(q)}{\sum_{q'} w_i^t(q')}.$$

Theoretical results on the convergence properties of reinforcement learning are scarce.[11] Thus most of the analysis is based on simulations. We ran several simulations of reinforce-

[10] We imposed a lower bound of 1 on propensities.

[11] Laslier, Topol, and Walliser (2001) show that reinforcement learning converges with positive probability to any strict pure Nash equilibrium in finite two-player strategic games. Similar results were obtained by Ianni (2002). However, they do not consider reinforcement of neighborhoods as in our case.

ment learning against itself as well as other decision rules while varying the initial propensities $w_i^1(q)$. Choosing initial propensity is always a bit arbitrary. However, results did not change much when using different initial propensities. We chose $w_i^1(q) = 78$, which minimized the mean squared deviation to the Nash equilibrium. Since reinforcement learning already is a stochastic process, we did not add additional noise to the process.

3.5 Trial & error (t&e)

Huck, Normann, and Oechssler (2004a) introduce a very simple trial & error learning process. Players begin by randomly adjusting their initial quantity either up- or downwards with an exogenously fixed step size. If this change increases profits, the direction is continued. If it does not, the direction of adjustment is reversed. We chose a step size of 4. Formally, players adjust their quantities as follows:

$$q_i^t := \max\{0, \min\{q_i^{t-1} + 4s_i^{t-1}, 109\}\},$$

for $t = 2, ...,$ where

$$s_i^t := \begin{cases} sign(q_i^t - q_i^{t-1}) \times sign(\pi_i^t - \pi_i^{t-1}) & \text{if} \quad (q_i^t - q_i^{t-1})(\pi_i^t - \pi_i^{t-1}) \neq 0 \\ +1, -1 \text{ each with positive probability otherwise.} \end{cases}.$$

On the boundaries of the output grid, we chose a "soft reflecting boundary". In particular, when the process reached 109 or 0 twice in subsequent periods, the next quantity chosen was $109 - 4$ or $0 + 4$, respectively.

Huck, Normann, and Oechssler (2004a) show that in Cournot duopoly if players choose the wrong direction with small but positive probability, then trial & error learning converges in the long run to a set of outcomes around the collusive outcome. To follow the theoretical setting, the noise for this process was modelled such that the computer chose the opposite direction from that prescribed by the theory with independent probability of 0.2 in each round.[12]

4 Experimental design

In total, 550 subjects participated in our internet experiment. Subjects played in a location of their own choice (home, office etc.), and at their own pace. Recruitment was done by

[12] Trial & error learning can be viewed as a special operationalization of learning direction theory by Selten and Buchta (1998). This theory assumes that players can judge in which direction better actions can be found. In the absence of information about demand and cost conditions, one interpretation is that the right direction can be found by determining which direction was successful last period.

email, internet newsgroups, and a University of Bonn student magazine. Each subject chose a nickname. Incentives were provided by publicly displaying a highscore after the experiment (like in computer games). We did not want to exclude (and implicitly select) subjects by technical means. In order to participate in our experiment, a standard web browser and a low-speed internet connection were sufficient.

Subjects could repeat the experiment as often as they desired,[13] either immediately or at some later time. Subjects were encouraged to repeat under the same user name as before.[14] While 550 subject played the first round ("first–timers"), we recorded 500 plays by "repeaters".

The sequence of events was as follows. After logging in, subjects were matched to a computer type. The computer type was displayed to subjects via a label (Greek letters) though subjects were not told how computer types were associated with labels. In the instructions (see Appendix A) subjects were told the following: "The other firm is always played by a computer program. The computer uses a fixed algorithm to calculate its output which may depend on a number of things but it cannot observe your output from the current round before making its decision."

A page with instructions was displayed to subjects. At any time during the experiment, subjects were able read the instructions and an example for calculating profits by opening a separate window on their computer. After reading the instructions, subjects could input their quantity for the first round. The computer displayed a new window with the results for the current round including the number of the round, the subject's quantity, the subject's profit, the computer's quantity as well as the computer's profit (see Appendix B for screenshots). Subjects had to acknowledge this information before moving on to the following round. Upon acknowledgment, a new page appeared with an input field for the new quantity. This page also showed a table with the entire history of previous rounds' quantities and profits for both players.[15]

After round 40, subjects were asked to fill in a brief questionnaire (see Appendix) with information on gender, occupation, country of origin, formal training in game theory or economic theory, previous participation in online experiments, and the free format question

[13] One subject actually played a record 31 times.

[14] The incentives for doing so were the highscore and the possibility to pick the same computer opponent as before (subjects logging in under a different name were allocated to a randomly chosen computer). The latter possibility was only revealed once subjects logged in under the same name.

[15] See the working paper version (Duersch et al., 2005) for an additional treatment, in which subjects were reminded only of the previous round's results. The results did not differ significantly.

"Please explain in a few words how you made your decisions". It was possible to skip this questionnaire. The highscore was displayed on the following page. This table contained a ranking among all previous subjects, separately for subjects who were matched against the same computer type and for all subjects. It also contained the computer's highscore.

5 Results

To give a first impression of the data, we present in Table 2 mean quantities of subjects and computers, respectively, averaged over all rounds and subjects. The first thing to notice is that subjects on average have much higher quantities than computers (48.68 vs. 33.29). This holds for all treatments except for the imitation treatment. Recall that the Cournot–Nash quantity is 36 (see Table 1). Thus, subjects chose on average quantities that exceeded by far the Cournot quantity and in some cases came close to the Stackelberg leader output of 54.

Table 2: Mean quantities

treatment	subjects' mean quantities	computers' mean quantities
br	51.99 (0.61)	27.79 (0.30)
t&e	48.96 (0.71)	32.05 (0.49)
fic	46.11 (0.74)	31.94 (0.26)
imi	46.40 (0.91)	48.38 (0.49)
re	47.45 (0.83)	35.71 (0.72)
Total	48.68 (0.34)	33.92 (0.29)

Note: Average quantities over all 40 rounds and all subjects in a given treatment. The Cournot-Nash equilibrium quantity is 36. Standard errors of means in parentheses.

5.1 How are profits affected by the opponent's learning algorithm?

How do subjects' profits differ with respect to their computer opponents? Figure 1 shows a boxplot, which compactly summarizes the range of subjects' average profits per round of first time players and repeaters, respectively. The figure reports those measures separately for each treatment, i.e. for each computer opponent (br, t&e, fic, imi, and re). In the boxplot, the boxes denote the interquartile range (ICR) between 75th and 25th percentiles, i.e. 50% of observations are concentrated in this range. The line in the box denotes the median profit. The length of the whiskers is the min of 1.5 times the ICR and the distance

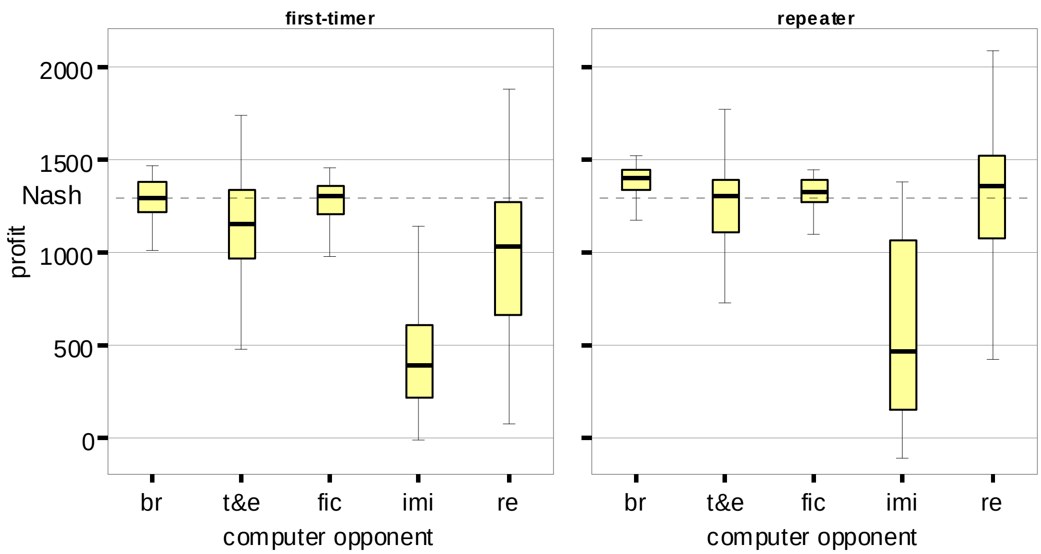


Figure 1: Boxplot of human subjects' profits against the different computer opponents (left panel: first–timers, right panel: repeaters).
The boxes denote the interquartile range (ICR) between 75th and 25th percentiles. The line in the box denotes the median profit. The length of the whiskers is the min of 1.5 times the ICR and the distance to the most extreme outcome. The dotted line shows the profit in the static Nash equilibrium.

to the most extreme outcome. The dotted line shows the profit per round in the Cournot Nash equilibrium for comparison.

First time players who are matched against computer types br, t&e, or fic achieve median profits that are about equal to or slightly less than the Nash equilibrium profit. Drastically different, however, are profits of subjects who were matched against the computer types imi and re. Median profits against imi were less than half the profits against the first three computer types. Even the very best subjects do not reach the Nash equilibrium profit, despite the bias in the noise of this computer type (see Footnote 7). Profits against computer type re are also substantially lower than against br and fic but they are higher than against imi.[16] The range of profits is highest against this type of computer. Some subjects achieve very high profits that exceed the Stackelberg leader or collusive profit (of 1458).

Median profits of repeaters are generally higher than those of first time players.[17] While subjects improve somewhat against computer type imi, median profits are still by far the lowest of all computer types. Against br, fic, and re subjects achieve higher median profits than in the Nash equilibrium. Again, the very best subjects played against t&e and re.

It is also quite instructive to consider average profits over time. Figure 2 shows average profits of subjects and computers for all 40 periods. Subjects playing against type br almost immediately gain a substantive edge over the computer and keep their profits more or less constant somewhere between the Stackelberg leader profit and the Nash equilibrium profit. The final result against type fic is similar but convergence is much more gradual. This shows a considerable amount of foresight on the side of our subjects. When playing against fic (in contrast to br), subjects must be more patient and forward looking to "teach" the computer into a Stackelberg follower position. The fictitious play computer is also the most successful among the computer types as it stabilizes at a profit of above 1000. The time series against types t&e and re look similar, although against the latter subjects do not even manage to achieve the Nash equilibrium profit on average.[18]

[16] For first-time players, profits against re are lower than against br and fic according to two–sided, non–parametric Mann–Whitney U tests (see e.g. Siegel and Castellan, 1988) at $p < 0.001$. For repeaters these differences are not significant anymore. For both, first-timers and repeaters, profits against imi are lower than against any other computer type at $p < 0.001$. A robust rank order test yields the same significance levels.

[17] Profits of those first–timers who we could identify as subsequent repeaters were actually lower in their first play than those of other first–timers although this difference is not significant. Thus, the increase in profits shown in Figure 1 appears to be driven by experience rather than selection of subjects.

[18] The dip of the computers' profits in round 2 is due to the high relative weight of the (uniformly

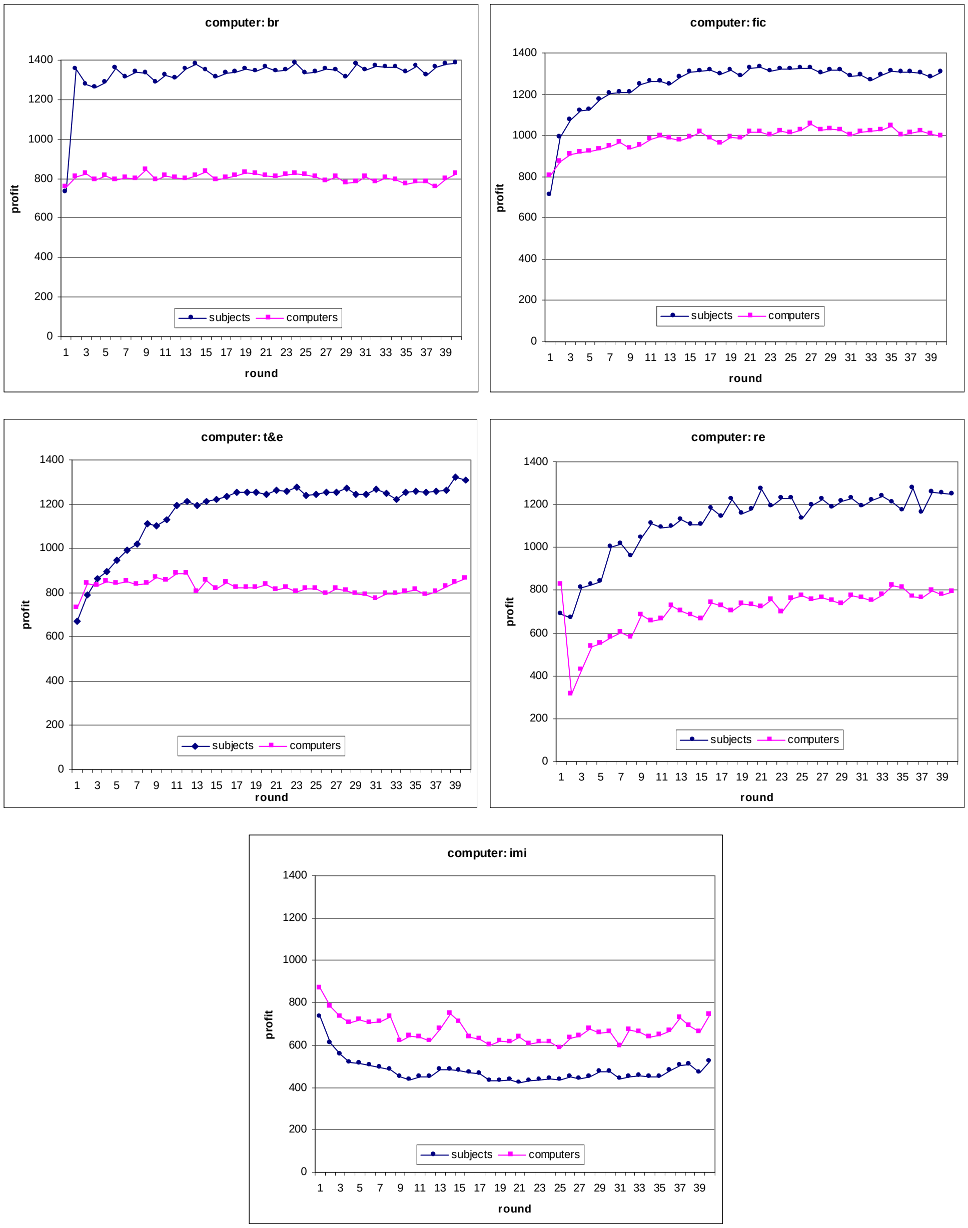


Figure 2: Time series of profits for subjects and computers for different computer types.

Computer type imi yields a totally different picture. In contrast to all others, it is the only computer type where subjects' payoffs are lower than those of computers. Furthermore, subjects average profits are much lower than against other learning algorithms. Finally, the data show that payoffs against imi decrease over time, both for subjects and for computers. All three results are in line with the theoretical predictions of Proposition 1.

Table 3 considers the subjects with the profits overall. Among the top 100 subjects, there are 48 subjects who played against a computer of type re, 28 who played against type br, and 24 who played against t&e. The top 10 players were almost exclusively playing against type re. This confirms the impression obtained from Figure 1. The highest profits can be achieved against type re but a lot of luck is needed for this due to the stochastic nature of reinforcement learning.

Table 3: Distribution of top subjects

against computer type...	among top 100	among top 10
br	28	–
t&e	24	1
re	48	9

Note: Pooled over first-timers and repeaters.

5.2 Human tactics

In this section we shall answer the second question raised in the introduction, namely how can the opponent's learning process be influenced by the player's behavior. One particularly intriguing question concerns evidence for strategic teaching, i.e. strategic manipulation of the computer types by subjects. Subjects may forgo short-term gains in order to manipulate the programmed opponent and earn large gains in later periods.[19]

Since initially, first–timers did not know the computer type, they may experiment with different quantities in order to explore the programmed opponent's responses. Table 4 reports the fraction of subjects that experiment with quantities. We call a subject ex-

distributed) initial weights in early rounds, while the computer quantity in round 1 is not chosen by the learning theory, but set to 35, 40 or 45.

[19] Collusion as an outcome is theoretically possible only against computer type t&e (Huck, Normann, and Oechssler, 2004a). However, as data on individual plays reveal, there were no successful examples of collusion between subject and computer over a prolonged period.

perimenting if the standard deviation of his quantities in the first 10 rounds is at least twice the standard deviation in the last 30 rounds. Overall at least one quarter of the first-time subjects experiment with quantities in this sense. The fraction exceeds 40% for fictitious play and trial & error learning. Note that these two are the "slowest moving" computer types. Table 4 also reports the fraction of repeaters who experiment. Interestingly, exploration declines when subjects play repeatedly except for reinforcement learning. So for all learning theories except reinforcement learning, exploration of first-timers may yield information that is used when the experiment is repeated. There may be two reasons for why it is different for reinforcement learning. First, note that reinforcement learning involves a probabilistic choice rule and may appear quite erratic to subjects. Therefore it may take more effort to learn about reinforcement learning than about other computer types. Second, as we have seen in previous sections, with some luck subjects can earn large profits if reinforcement starts with low quantities. Subjects' experimentation in the first 10 rounds may be aimed exactly at this.

Table 4: Classification of tactics

against computer type...	tactic	first-timer	repeater
br	Experimentation	25%	18%
	Leadership	20%	32%
fic	Experimentation	41%	29%
	Leadership	14%	11%
re	Experimentation	25%	32%
	Leadership	16%	16%
imi	Experimentation	31%	25%
	Leadership	16%	15%
t&e	Experimentation	40%	18%
	Leadership	16%	28%

Once subjects have explored and learned about the computer type, they may use this information to actively manipulate the computer type. Such manipulations may take on various forms. Probably the most straightforward form of manipulation is aimed at achieving Stackelberg leadership through aggressive play of large quantities. Table 4 also reports the fraction of subjects with such leadership behavior. We define a subject as displaying leadership behavior if he chooses a quantity of least 50 for at least 36 out of 40 rounds. About 16% of the first-timers display such leadership behavior. When playing against best response or trial & error learning, this behavior becomes even more pronounced among repeaters. The increase in leadership behavior is most remarkable when subjects play against br. Indeed, playing aggressively is a quite successful manipulation of br. Figure 3(a) shows quantities of the most successful subject playing against br and the corresponding computer quantities. This subject (ranked overall 45^{th}) chose 55 in all 40 periods.[20] The computer quickly adjusted to a neighborhood of the Stackelberg follower quantity with the remaining movement due to the noise in the computer's decision rule.

While leadership may be a relatively simple form of strategic manipulation, individual data reveal manipulations that can be very sophisticated. We discovered quite interesting, though not very frequent, patterns that can be seen in Figure 3(b). The subject who played against best response chose – with only slight variations – the following cycle of 4 quantities: 108, 70, 54, 42, 108, 70, ... Stunningly, this cycle produces an expected profit per round of 1520, which exceeds the Stackelberg leader profit.[21] By flooding the market with a quantity of 108, the subject made sure that the computer left the market in the next period. But instead of going for the monopoly profit, the subject accumulated intermediate profits over three periods. This, of course, raises the question, whether a cycle is optimal and how the optimal cycle looks like. It turns out, that in this game a cycle of length is four is optimal and, after rounding to integers, the optimal cycle is 108, 68, 54, 41, which produces an expected profit of 1522.[22] Thus, our subject was within 2 units of the solution for this non–trivial optimization problem.[23]

How did the very best subject play? Like all top players, he played against computer type re. Figure 3(c) reveals that the subject simply got lucky. The reinforcement algorithm locked in at very low quantities in the range between 10 and 20, and the subject roughly played a best response to that, which resulted in an average profit of 2091.

One benchmark to compare the behavior of our subjects to is the maximal profit an omniscient, myopic player could achieve against the respective learning theory. To generate this benchmark, we ran simulations pitting our 5 computer types against a simulated player who can perfectly forecast the action his computer opponent is about to take (including the noise) and plays a best response to that, but disregards the influence of his action on the future behavior of his opponent. As Figure 4 shows, our repeater subjects outperform that

[20] Curiously, none of our subjects chose the exact Stackelberg leader quantity of 54.

[21] The only reason the subject in Figure 3(a) received an even higher payoff was luck due to favorable noise of the computer algorithm.

[22] See Schipper (2006) for a proof of this claim.

[23] The subject played three times against br and left two comments. The first was "tried to trick him", the second "tricked him".

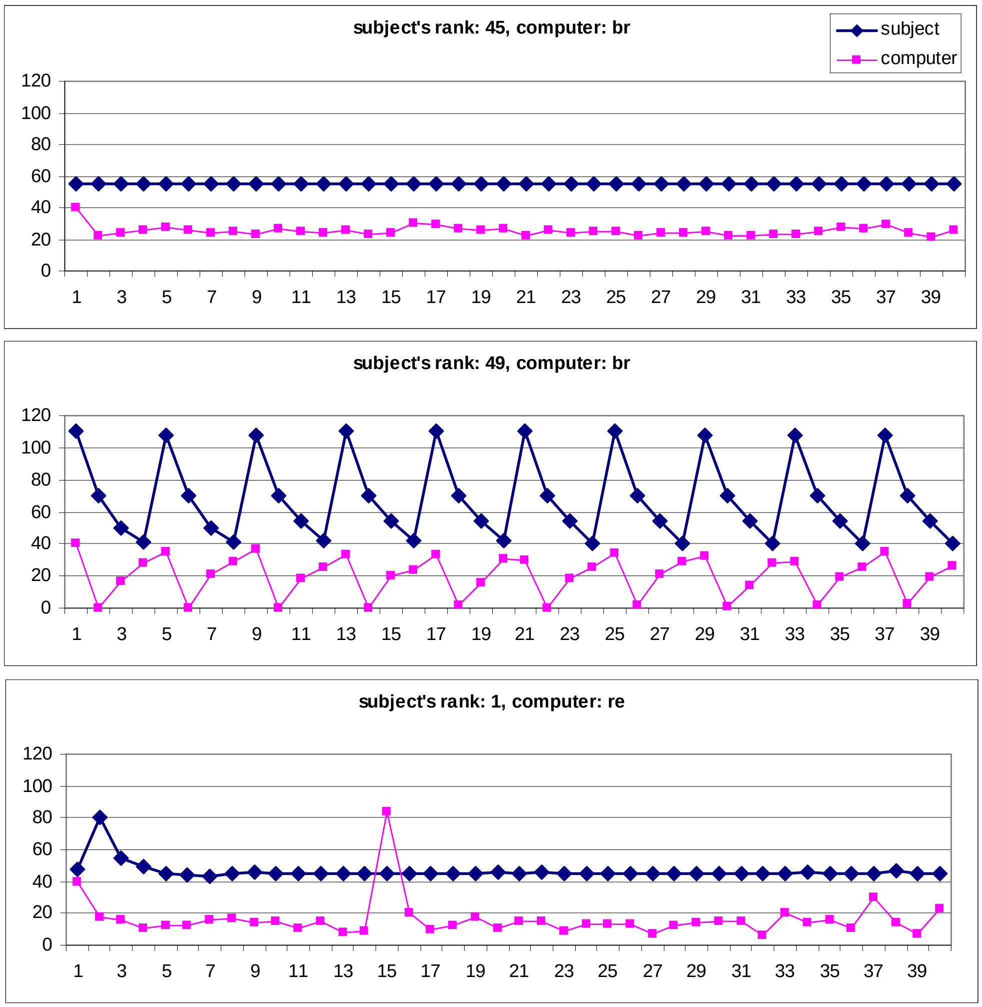


Figure 3: (a) Quantities of subject ranked number 45 and of the br-computer opponent (top panel); (b) Quantities of subject ranked number 49 and of the br-computer opponent (middle panel); (c) Quantities of top-ranked subject and his re-computer opponent

benchmark against br, re, and t&e. They do worse than the benchmark against fictitious play but considerably worse only against imitate the best. Given that the myopic best behavior requires a huge amount of knowledge about the opponent, which our subjects can not possibly possess since each learning theory incorporates a random element, the only way for subjects to match or outperform the best myopic benchmark is by playing more sophisticated than myopic – by influencing future play of the learning theories.

6 The role of incentives: a robustness check

In the main (internet) experiment, the only incentive for subject was the highscore. It is a justified question whether these incentives are strong enough compared to the usual financial incentives given in laboratory experiments. To check for this, we conducted a control experiment as a regular laboratory experiment with the usual monetary incentives. In the lab experiment, 50 subjects played in the Bonn Laboratory for Experimental Economics. The instructions and the computer interface for both settings were the same up to the incentive structure. Subjects were required to repeat the experiment once with the same computer type as opponent, i.e., they played two times 40 rounds. Since there were fewer observations in the lab, we used only a starting value of 40 for the computer types.[24] Incentives were provided by paying subjects immediately at the end of the experiment the sum of profits over all rounds according to an exchange rate of 9000 Points to 1 euro. On average, subjects earned 10.17 euro for about half an hour in the lab.

We do find some significant differences between the two incentive structures. In the lab experiment, average quantities are significantly lower (MWU–test, $p < 0.01$) although average profits do not significantly differ.[25] However, the crucial point is whether the differences across our treatments are robust to changing the incentive structure. The left panel of Figure 5 shows average profits of all subjects given their respective computer opponent. The only significant difference between the lab and the internet is for computer opponent imi (MWU–test, $p < 0.05$). More importantly, all treatment effects are qualitatively the same, independently of the incentive structure.[26] The right panel of Figure 5 shows the

[24] Recall that in the internet experiment, computer algorithms had an equal chance of starting with values 35, 40 or 45.

[25] Both results hold for all subjects and for first–timers only. All tests in this Section can also be conducted as Kolmogorov-Smirnov tests without changing the results. For additional tables with data for each learning algorithm see the working paper version (Duersch, Kolb, Oechssler, and Schipper, 2008).

[26] Likewise, when we recalculate Tables 3 and 4 by including data from the lab, only very minor changes occur.

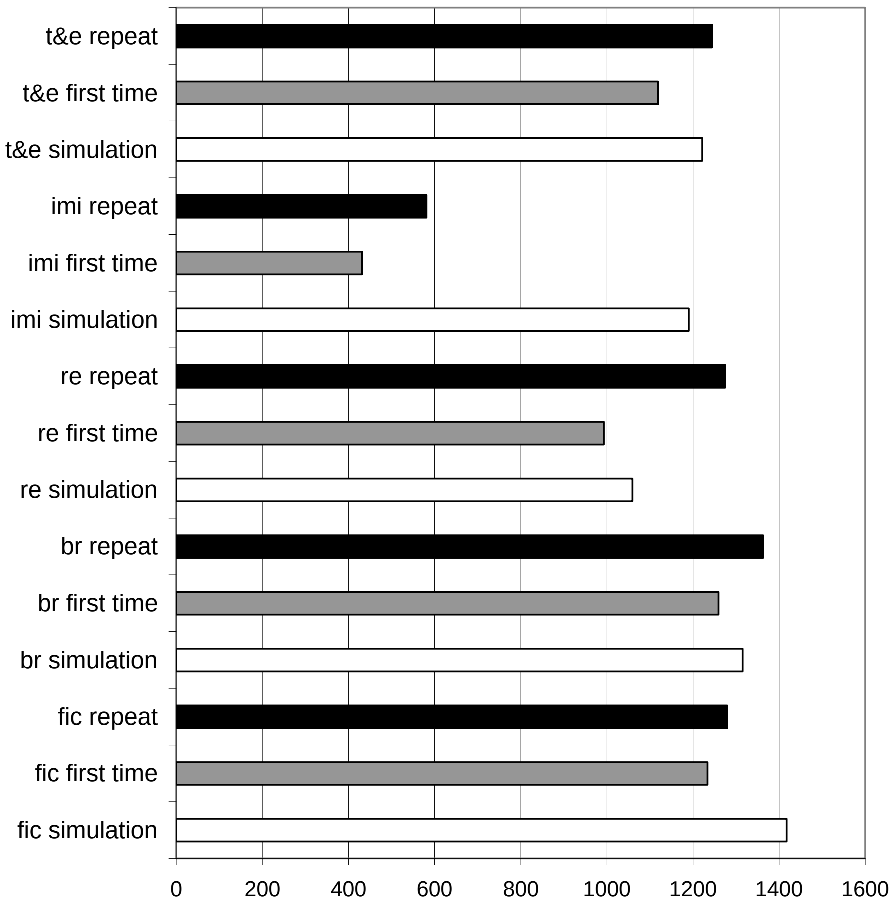


Figure 4: Average profits per round of simulated omniscient, myopic player (white bars) vs. actual profits of repeaters (black bars) and first-time subjects (grey bars) when matched against the different computer types (e.g. re repeat is the average profit of repeaters against computer re, re simulation is average profit of the omniscient player against re. Note: The omniscient player can perfectly predict the computer's action (including noise).

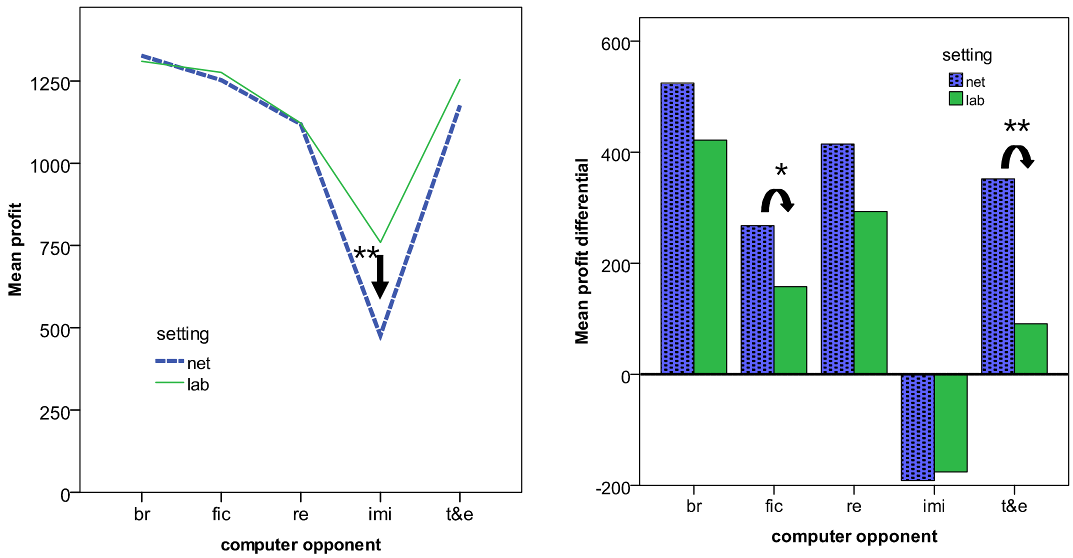


Figure 5: A comparison of the internet experiment (net) and the laboratory (lab). (a) Left panel: Mean profits of human subjects against their respective computer opponents. (b) Right panel: Mean difference between the subject's profit and the computer's profit.
Note: ** significant difference at the 1% level ; * significant difference at 5% level, MWU–tests.

average difference of the subject's profit and the computer's profit. The main result, that human subjects manage to exploit all of their computer opponents except imi, holds for both incentive structures. The average profit differential is significantly larger than 0 for computer opponents br, fic, re, and t&e and for both incentive structures, net and lab.[27] The profit differential against imi is significantly smaller than 0 for both incentive structures at the 1% level. Furthermore, testing across incentive structures, profit differentials against br, re, and imi are not significantly different between net and lab. Profit differentials against fic and t&e are significantly larger in net but, qualitatively, the treatment effects seem to be robust.

[27] According to t–tests at the 1% level with the exception of t&e in lab for which the differential is only marginally larger at the 10% level. Results for a Wilcoxon test are qualitatively the same.

7 Conclusion

In this experiment we let subjects play against computers which were programmed to follow one out of a set of popular learning theories. The aim was to find out whether subjects were able to exploit those learning algorithms. We find that there are remarkable differences in the exploitability of myopic learning algorithms. There are two insights from this observation: First, while the bulk of traditional learning theories that have been studied in the literature are myopic in nature, we need more advanced learning theories that incorporate at least a limited amount of foresight if those learning theories should also fit subjects engaged in strategic teaching. Many of our subjects were quite able to exploit the simple myopic learning algorithms. Strategic teaching is an important phenomenon that needs to be accounted for in the future development of theory.

Second, from an evolutionary perspective, one desideratum to impose on a realistic learning theory should be its non-exploitability. If a learning theory can be exploited by sophisticated subjects to their advantage, then such a learning theory should disappear in the long run. Interestingly, we find that among the learning algorithms studied in this paper, only imitate-the-best is robust to exploitation. This learning algorithm is known to lead to a non-Nash equilibrium in the game studied. This observation poses the following theoretical question left for further research: Does there exist a simple adaptive learning algorithm that is not exploitable and leads to Nash equilibrium?

References

[1] Alós-Ferrer, C., Ania, A.B.: The evolutionary stability of perfectly competitive behavior, Econ Theory, 26, 497-516 (2005)

[2] Apesteguia, J., Huck, S., Oechssler, J.: Imitation - Theory and experimental evidence, J Econ Theory, 136, 217-235 (2007)

[3] Brown, G.W.: Iterative solutions of games by fictitious play, in: Koopmans, T.C. (ed.), Activity analysis of production and allocation, John Wiley (1951)

[4] Camerer, C., Ho, T.H.: Strategic learning and teaching in games, in: Hoch, S., Kunreuther, H. (eds.) Wharton on decision making, New York: Wiley (2001)

[5] Camerer, C., Ho, T.H., Chong, J.K.: Sophisticated experience-weighted attraction Learning and strategic teaching in repeated games, J Econ Theory, 104, 137-188 (2002)

[6] Cason, T., Sharma, T.: Recommended play and correlated equilibria: An experimental study, Econ Theory, 33, 11-27 (2007)

[7] Coricelli, G.: Strategic interaction in iterated zero-sum games, mimeo (2005)

[8] Cournot, A.: Researches into the mathematical principles of the theory of wealth, transl. by N. T. Bacon, MacMillan Company, New York, 1927 (1838)

[9] Drehmann, M., Oechssler, J., Roider, A.: Herding and contrarian behavior in financial markets, Am Econ Rev, 95(5), 1403-1426 (2005)

[10] Duersch, P., Kolb, A., Oechssler, J., Schipper, B.: Rage Against the Machines - How Subjects Learn to Play Against Computers, AWI–Discussion Paper No. 423, Department of Economics, University of Heidelberg (2008)

[11] Duersch, P., Oechssler, J., Schipper, B.: Experimenting on the internet: Does it make a difference?, mimeo, University of Heidelberg (2008)

[12] Ellison, G.: Learning from personal experience: One rational guy and the justification of myopia, Game Econ Behav, 19, 180-210 (1997)

[13] Erev, I., Roth, A.: Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria, Am Econ Rev, 88, 848-881 (1998)

[14] Erev, I., Haruvy, E.: "Learning and the Economics of Small Decisions", forthcoming in Kagel, J.H., Roth, A.E. (eds.) The Handbook of Experimental Economics, Vol. 2 (2008)

[15] Fudenberg, D., Levine, D.K.: The theory of learning in games, Cambridge: MIT Press (1998)

[16] Fudenberg, D., Levine, D.K.: Reputation and equilibrium selection in games with a patient player, Econometrica, 57, 759-778 (1989)

[17] Huck, S., Normann, H.T., Oechssler, J.: Learning in Cournot oligopoly: An experiment, Econ J, 109, C80-C95 (1999)

[18] Huck, S., Normann, H.T., Oechssler, J.: Through trial & error to collusion, Int Econ Rev, 45, 205-224 (2004)

[19] Huck, S., Normann, H.T., Oechssler, J.: Two are few and four are many: Number effects in experimental oligopoly, J Econ Behav Organ, 53, 435-446 (2004)

[20] Ianni, A.: Reinforcement learning and the power law of practice: Some analytical results, University of Southampton (2002)

[21] Laslier, J.-F., Topol, R., Walliser, B.: A behavioral learning process in games, Game Econ Behav, 37, 340-366 (2001)

[22] Monderer, D.,Shapley, L.: Potential games, Game Econ Behav, 14, 124-143 (1996)

[23] Offerman, T., Potters, J., Sonnemans, J.: Imitation and belief learning in an oligopoly experiment, Rev Econ Studies, 69, 973-997 (2002)

[24] Possajennikov, A.: Evolutionary foundation of aggregative-taking behavior, Econ Theory, 21, 921-928 (2003)

[25] Robinson, J.: An iterative method of solving games, Ann Math, 54, 296-301 (1951)

[26] Roth, A., Erev, I.: Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term, Game Econ Behav 8, 164-212 (1995)

[27] Sarin, R., Vahid, F.: Strategic similarity and coordination, Econ J, 114, 506-527 (2004)

[28] Schipper, B.C.: Strategic control of myopic best reply in repeated games, University of California, Davis (2006)

[29] Schipper, B.C.: Imitators and optimizers in Cournot oligopoly, University of California, Davis (2008)

[30] Schipper, B.C.: Submodularity and the evolution of Walrasian behavior, Int J Game Theory, 32, 471-477 (2003)

[31] Selten, R., Buchta, J.: Experimental sealed bid first price auctions with directly observed bid functions, in: Budescu, D., Erev, I., Zwick, R. (eds.), Games and human behavior: Essays in honor of Amnon Rapoport, Lawrence Erlbaum Associates, Mahwah, N.J. (1998)

[32] Shachat, J., Swarthout, J. T.: Learning about learning in games through experimental control of strategic independence, University of Arizona (2002)

[33] Siegel, S., Castellan, N.J., Jr.: Nonparametric Statistics for the Behavioral Sciences, Singapore: McGraw-Hill (1988)

[34] Tanaka, Y.: Long run equilibria in an asymmetric oligopoly, Econ Theory, 14, 705-715 (1999)

[35] Vega–Redondo, F.: The evolution of Walrasian behavior, Econometrica, 65, 375-384 (1997)

Appendix

A Proof of Proposition 1

Proof. Let the computer be player 2. We first claim that imi will imitate player 1's quantity from the previous round if and only if it is closer to 54 than its own quantity and $q_1+q_2 \neq 108$. For $q_1+q_2=108$, both profits are zero and imi sticks to its own strategy. For $q_1+q_2 \neq 108$, we need to show that $\pi(q_1,q_2) > \pi(q_2,q_1)$ if and only if $|q_1-54| < |q_2-54|$. For $q_1+q_2 < 108$ or for $q_1, q_2 > 54$, the claim follows because for $p > c$ ($p < c$) the firm with higher (lower) quantity makes the larger profit (the smaller loss). It remains to consider the case $q_i > 54 > q_{-i}$ but $q_1+q_2 > 108$. In this case

$$|q_i - 54| = q_i - 54 > 54 - q_{-i} = |q_{-i} - 54| \, .$$

Since $p < c$, both firms make losses and since $q_i > q_{-i}$, firm i's losses are larger, which proves the claim. Note in particular, that as long as $q_1+q_2 < 108$, imi will never lower its quantity q_2.

(a) Since quantities are strategic substitutes, the above claim implies that for starting values of the computer higher than the Cournot quantity, $q_2^1 > 36$, player 1 is best off by not triggering imi to choose higher quantities than q_2^1. Thus, imi will continue to play q_2^1 and player 1's optimal strategy is to play a myopic best reply $q_1 \in br(q_2^1)$ every period. The resulting profits for the human player over 40 rounds are $\max_{q_1} \pi(q_1, 40) = 46\,240$ and $\max_{q_1} \pi(q_1, 45) = 39\,680$. The resulting profits for computer imi are $\pi(40, q_1) = 54\,440$ and $\pi(45, q_1) = 57\,600$.

If $q_2^1 = 35$, player 1's optimal strategy is to play $q_1^t = q_2^1$, for all $t = 1, ..., 39$ and $q_1^{40} \in br(q_2^1)$. This results in a profit of 53 202. The imi computer's profit is 53 165 in this case. Since the three starting values 35, 40, and 45 occur with the same frequency, the average profit for the player from this optimal strategy is 46 374, while his computer opponent imi earns 55068.

(b) Given a starting value q_2^1 of the computer, the player can obtain a positive profit differential in a given round t if and only if $\left|q_1^t - 54\right| < \left|q_2^t - 54\right|$. Unless the player changes his strategy again, this profit differential is eroded in the next period due to the adjustment of imi to the quantity that yielded the higher profit. The player can close the gap to 54 in one large step or in several small steps. Due to the linear structure of the Cournot game, the number of steps does not matter. To see this note that the profit differential for one step is given by $(p-1)(q_1 - q_2^1)$. For starting values of imi below 54, this expression is maximized by $q_1 = 54$. Now suppose the player takes two arbitrary steps in rounds 1 and 2 to reach 54 (for all rounds thereafter the profit differential is zero). The profit differential is then given by

$$(p^1 - 1)(q_1^1 - q_2^1) + (p^2 - 1)(q_1^2 - q_2^2). \tag{5}$$

Since imi always imitates quantities that are closer to 54, we have that $q_2^2 = q_1^1$. Thus, (5) is maximized by $q_1^2 = 54$ and arbitrary q_1^1 such that $q_2^1 \leq q_1^1 \leq 54$. Consequently, the profit differential is the same with one or two steps. The same argument holds for any number of steps towards 54.

The maximal profit differentials that can be obtained for the three starting values 35, 40, and 45 are therefore 361, 196, and 81, respectively, which yield an average profit differential of 212.67. ■

B Instructions

B.1 Introduction Page

Welcome to our experiment!

Please take your time to read this short introduction. The experiment lasts for 40 rounds. At the end, there is a high score showing the rankings of all participants. You represent a firm which produces and sells a certain product. There is one other firm that produces and sells the same product. You must decide how much to produce in each round. The capacity of your factory allows you to produce between 0 and 110 units each round. Production costs are 1 per unit. The price you obtain for each sold unit may vary between 0 and 109 and is determined as follows. The higher the combined output of you and the other firm, the lower the price. To be precise, the price falls by 1 for each additional unit supplied. The profit you make per unit equals the price minus production cost of 1. Note that you make a loss if the price is 0. Your profit in a given round equals the profit per unit times your output, i.e. profit = (price - 1) * Your output. Please look for an example here. At the beginning of each round, all prior decisions and profits are shown. The other firm is always played by a computer program. The computer uses a fixed algorithm to calculate its output which may depend on a number of things but it cannot observe your output from the current round before making its decision. Your profits from all 40 rounds will be added up to calculate your high score. There is an overall high score and a separate one for each type of computer. Please do not use the browser buttons (back, forward) during the game, and do not click twice on the go button, it may take a short while.

Choose new quantity

Please choose an integer (whole number) between 0 and 110.

B.2 Example Page

The Formula

The profit in each round is calculated according to the following formula:

Profit = (Price − 1) * Your Output

The price, in turn, is calculated as follows.

Price = 109 − Combined Output

That is, if either you or the computer raises the output by 1, the price falls by 1 for both of you. (but note that the price cannot become negative). And the combined output is simply:

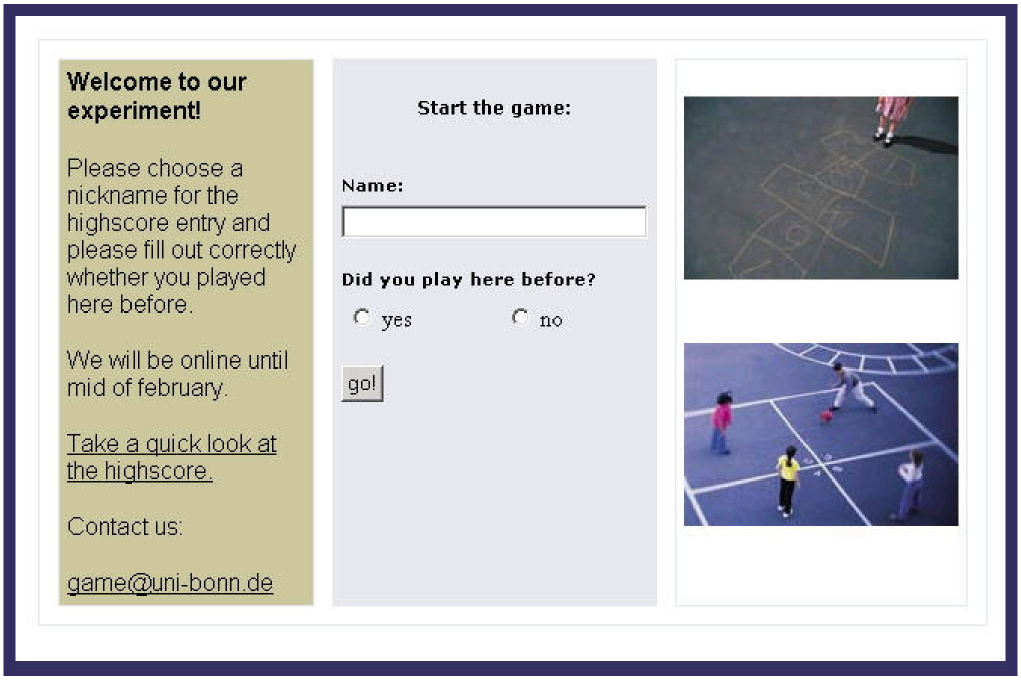
Combined Output = Your Output + Computers Output

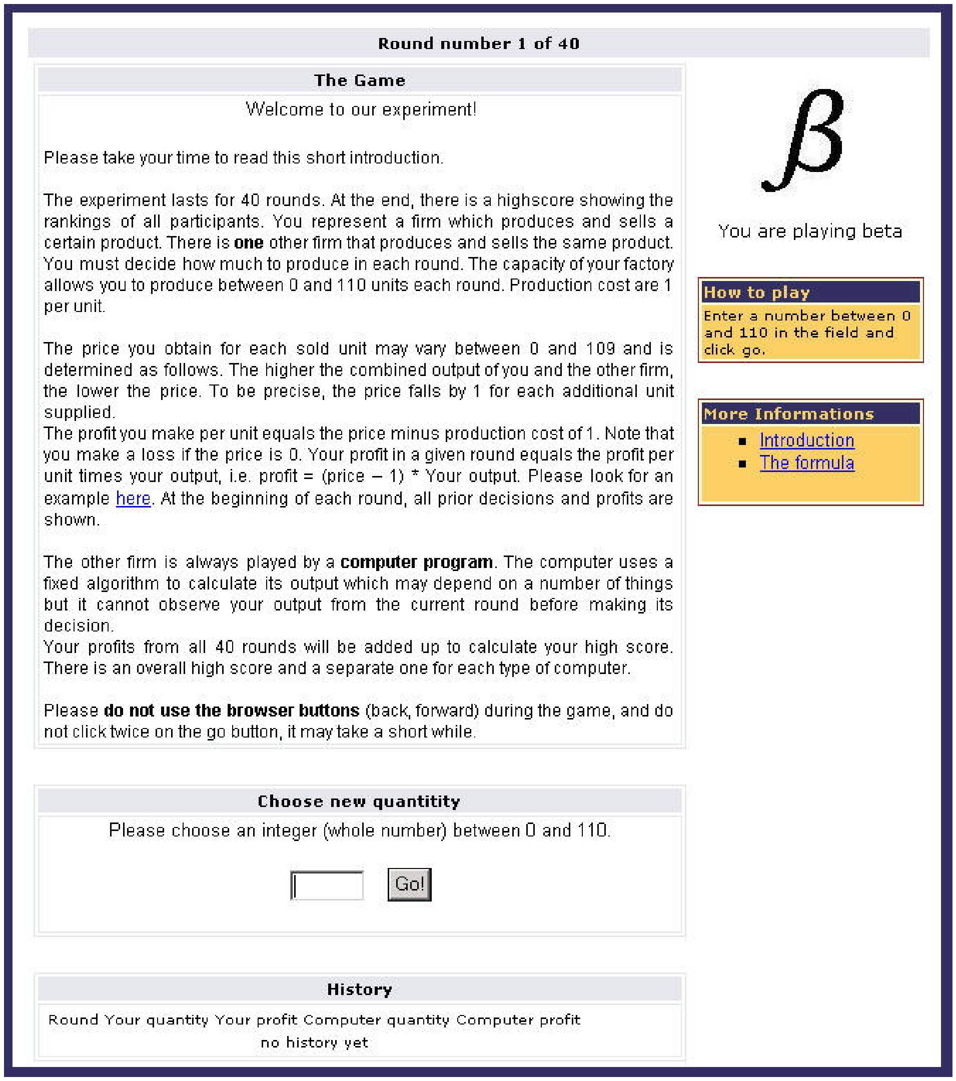
Example:

Lets say your output is 20, and the computers output is 40. Hence, combined output is 60 and the price would be 49 $(= 109 - 60)$. Your profit would be $(49 - 1) * 20 = 960$. The computers profit would be $(49 - 1) * 40 = 1920$. Now assume you raise your output to 30, while the computer stays at 40. The new price would be 39 $(= 109 - 40 - 30)$. Your profit would be $(39 - 1) * 30 = 1140$. The computers profit would be $(39 - 1) * 40 = 1520$.

To continue, please close this window.

C Screenshots





Round number 1 of 40

The Game

Welcome to our experiment!

Please take your time to read this short introduction.

The experiment lasts for 40 rounds. At the end, there is a highscore showing the rankings of all participants. You represent a firm which produces and sells a certain product. There is **one** other firm that produces and sells the same product. You must decide how much to produce in each round. The capacity of your factory allows you to produce between 0 and 110 units each round. Production cost are 1 per unit.

The price you obtain for each sold unit may vary between 0 and 109 and is determined as follows. The higher the combined output of you and the other firm, the lower the price. To be precise, the price falls by 1 for each additional unit supplied.
The profit you make per unit equals the price minus production cost of 1. Note that you make a loss if the price is 0. Your profit in a given round equals the profit per unit times your output, i.e. profit = (price – 1) * Your output. Please look for an example here. At the beginning of each round, all prior decisions and profits are shown.

The other firm is always played by a **computer program**. The computer uses a fixed algorithm to calculate its output which may depend on a number of things but it cannot observe your output from the current round before making its decision.
Your profits from all 40 rounds will be added up to calculate your high score. There is an overall high score and a separate one for each type of computer.

Please **do not use the browser buttons** (back, forward) during the game, and do not click twice on the go button, it may take a short while.

β

You are playing beta

How to play

Enter a number between 0 and 110 in the field and click go.

More Informations

- Introduction
- The formula

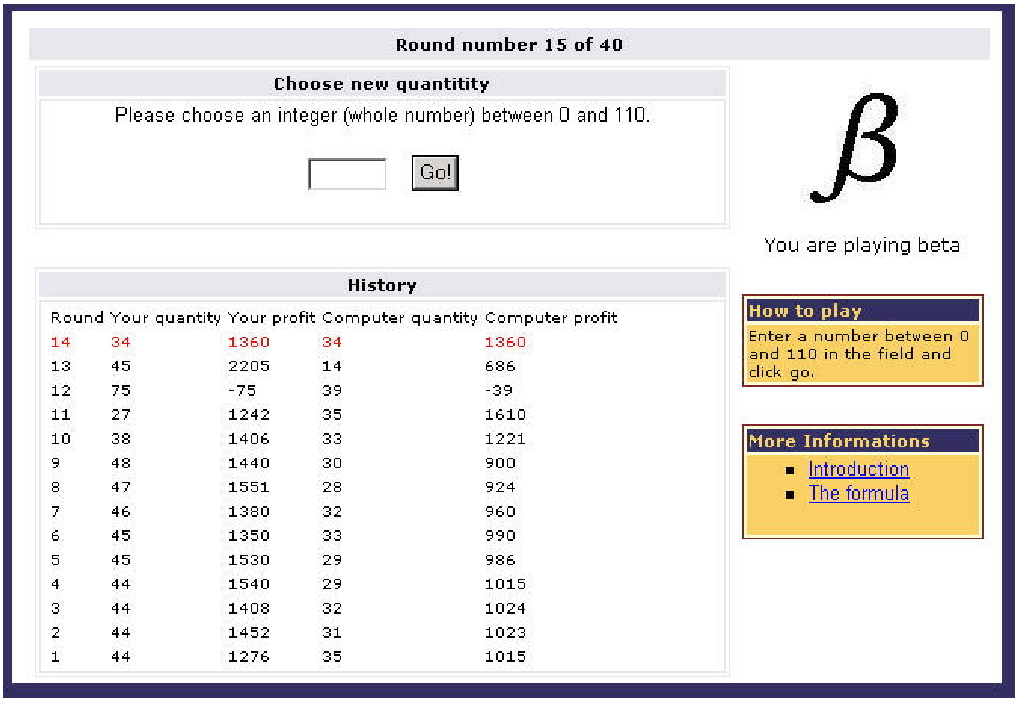
Choose new quantitity

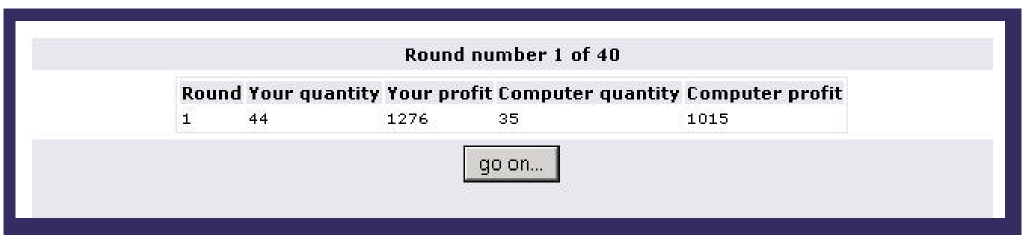
Please choose an integer (whole number) between 0 and 110.

Go!

History

Round Your quantity Your profit Computer quantity Computer profit
no history yet


Round number 15 of 40
Choose new quantitity
Please choose an integer (whole number) between 0 and 110.
Go!
You are playing beta
History
Round Your quantity Your profit Computer quantity Computer profit
14 34 1360 34 1360
13 45 2205 14 686
12 75 -75 39 -39
11 27 1242 35 1610
10 38 1406 33 1221
9 48 1440 30 900
8 47 1551 28 924
7 46 1380 32 960
6 45 1350 33 990
5 45 1530 29 986
4 44 1540 29 1015
3 44 1408 32 1024
2 44 1452 31 1023
1 44 1276 35 1015
How to play
Enter a number between 0 and 110 in the field and click go.
More Informations
Introduction
The formula


Round number 1 of 40
Round Your quantity Your profit Computer quantity Computer profit
1 44 1276 35 1015
go on...

II. Taking Punishment Into Your Own Hands: An Experiment on the Motivation Underlying Punishment*

Abstract

In a punishment experiment, we separate the demand for punishment in general from a possible demand to conduct punishment personally. Subjects experience an unfair split of their earnings from a real effort task and have to decide on the punishment of the person who determines the distribution. First, it is established whether the allocator's payoff is reduced and, afterwards, subjects take part in a second price auction for the right to (physically) carry out the act of payoff reduction. This auction only resolves who will punish, not whether punishment takes place, so only subjects with a demand for personal punishment should bid.

*Paper co-authored by Julia Mueller.

If the person who had done us some great injury, who had murdered our father or our brother, for example, should soon afterwards die of a fever, or even be brought to the scaffold upon account of some other crime, though it might sooth our hatred, it would not fully gratify our resentment. Resentment would prompt us to desire, not only that he should be punished, but that he should be punished by our means, and upon account of that particular injury which he had done to us. (Adam Smith[1])

1 Introduction

The desire for revenge, to punish those who did wrong upon oneself, is a strong motivation for humans. From ancient Greek dramas to modern movies, hardly a storyline can do without. It has also been the focus of extensive research in economics, both in the form of experiments which find that, indeed, subjects are willing to forgo monetary gains to exert punishment, and in the form of theoretical models that seek to explain such behavior. However, both the quote by Adam Smith above and several of those movies[2] feature a very specific observation about punishment: According to Adam Smith, humans not only care about punishment being inflicted on the perpetrator of a crime against them, but they also value carrying out that punishment themselves, personally. It is this, personal, characteristic of punishment that we try to isolate in the laboratory. Our experiment is designed to exclude other possible reasons why one would be willing to give up money to punish, e.g. subjects do not have to spend money to assure punishment is carried out, they only spend money to assure it is carried out by them personally.

Punishment has been documented in various experiments, especially in social dilemma situations where individual and group incentives diverge and free-riding occurs. One of the first experiments of this kind was conducted by Ostrom et al. (1992), where subjects who played various rounds in a common pool resource game were willing to pay a fee to place a fine on other subjects who over-extracted the resource. Fehr and Gächter (2000) demonstrate that costly punishment of free-riders who do not contribute occurs in a public goods experiment, with punishment leading to higher levels of cooperation. Nikiforakis and Normann (2008) analyze the effectiveness of such peer-imposed punishment in a public good game, finding that contributions increase in the effectiveness. In contrast, Falkinger et al. (2000) use punishment imposed "automatically" by the experimenter on non-contributers. Both peer-imposed and experimenter-imposed

[1] In *The Theory of Moral Sentiments*, page 113.

[2] E.g. Marsellus telling Butch to move aside, so he can shoot Zed in Pulp Fiction, or Grace shooting Tom herself in Dogville, "Some things, you have to do yourself".

punishment raises contributions. However, subjects are not only motivated by the monetary consequences of punishment. As Masclet et al. (2003) show, even non-monetary punishment, the expression of disapproval by others, leads to the same result. Masclet et al. (2003) are mainly interested in the receiving side of the punishment, but, it is also interesting to investigate the decision process of the punishing side.

Direct neuroeconomic evidence that subjects "like" to punish was found by De Quervain et al. (2004). They used PET recordings of brain activation to investigate the mechanisms in the brain involved in punishment. Subjects played a trust game where cooperating players could punish defecting partners. In the punishment condition activation of the dorsal striatum was found, which is well known for its reward processing properties. This could either be due to the fact that the defecting partner lost money or it could be pleasure derived from the act of punishing. Based on their finding that subjects do not condition the amount of their own punishment onto the punishment already dealt (to the same person) by other subjects, Casari and Luini (2009), speculate that in the same vein that "the punisher derives her utility from the act of punishment in itself and not from achieving, in conjunction with other punishers, a total amount of punishment."

Spurred on by the experimental observation that people do not always act purely selfish, new theories of other-regarding preferences have been put forward, capturing phenomena like fairness, altruism, inequity aversion. Levine (1998) uses an adjusted utility which is supplemented by a term which takes into account the opponent's utility weighted by an altruism coefficient. Inequity aversion models add to the utility derived from own income a term that represents concern about the payoff distribution; Fehr and Schmidt (1999) use the difference between subjects own payoff and the payoffs of the opponents, Bolton and Ockenfels (2000) the proportion of total payoffs.

Other theories develop techniques to embed concerns for reciprocity. Rabin (1993) models reciprocity in normal form games by adding psychological payoffs to the material payoffs. This additional term captures intentions via beliefs of the players and defines the kindness of players in relation to his possible actions. Dufwenberg and Kirchsteiger (2004) dilate this approach to sequential games. Falk and Fischbacher (2006) also transform standard games into psychological games. In their model, utility of the players depends not only on the payoffs but also on a reciprocity term which embodies kindness and reciprocation.

All of these theories incorporate the opponent's outcome into the utility of the player, and several can explain reciprocal behavior or punishment. However, we are not primarily interested in the fact that the payoff of an offender is reduced, but especially in *who* will derive satisfaction from punishing. Only the person who conducted the punishment? Or everyone who saw the offender being punished, even if the punishment was not conducted "personally"?

Perhaps the theory closest to our design is the paper by Andreoni (1990). He examines private provision of a public good and models the utility of the individuals as a function not only of the amount of the public good but also of the own gift to the public good. This individual donation produces what Andreoni calls a "warm glow", utility derived from the act of giving. If one assumes in almost the same manner that the act of punishing enters the utility function, one would arrive at a theory that could account for a demand to punish personally.

In the next section, we introduce the design we use to investigate personal punishment. In section 3 we explain the theoretical solutions, then in section 4 our hypotheses and in section 5 we present our results. In section 6 we introduce our control experiment. Finally in section 7, we conclude with a discussion.

2 Experiment

2.1 Design

To test the demand for personal punishment, we use a design with four stages. Subjects are matched in groups of four; each group consists of three subjects A and one subject B. The experiment was anonymous, so no subject knew about the other subjects he or she was matched with.

Instructions for the experiment, which fully described the experiment for both type A and type B, were handed to subjects at the very start of the experiment, followed by test questions to check whether the subjects had understood the instructions.[3] Only when all subjects had correctly answered these questions, did the experiment proceed.

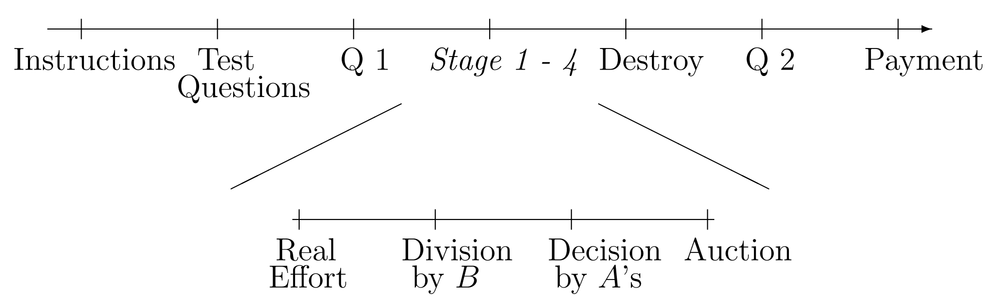


Figure 1: Timing

[3]See appendix A.1 and A.4 for translated instructions and test questions.

In the first stage, all subjects A participated in a real effort task where they could earn EUR 10. They were asked to fill a sheet of graph paper (A5, 148 × 210 mm, about 1260 squares) with alternating o and + signs. The allocated time frame was 25 minutes. Subjects who did not finish the task in time did drop out of the experiment and received no money apart from the show up fee. We chose this particular task for two reasons: First, it is simple and does not require any special abilities, so all subjects should be equally fit for the task. Second, as we found out in previous tests, the task is considerably more exhausting than it appears. We wanted to induce a feeling of ownership in those subjects who completed it. On the other hand, it was to look easy to the non-participating subjects B. During the task, all B's were sitting in the same room as the A's, but without any assignment.

After the task, the experimenters collected the sheets and informed each B how many A's in her group had succeeded. Upon learning that information, in stage two, B had to decide on an allocation of the money earned by the A's in the previous stage. The only two allocations available were (2,8): EUR 2 for A, EUR 8 for B, or (10,0): EUR 10 for A, EUR 0 for B. B could only implement the same allocation for all three A's she was matched with, not different allocations for different A's. So in the case of three successful A's subject B had to decide between EUR 24 for herself and EUR 2 for each A or EUR 0 for herself and EUR 10 for each A.

Before stage three, the experimenters informed all A's about the decision of their matched B. The money that B allocated to A was handed to A. The money that B allocated to herself was *also handed to* A, however it was put in an envelope by the experimenters. Then all A's had to decide whether they wanted to reduce B's payoff by destroying one of the three envelopes designated for B. If all A's decided not to reduce, the envelopes were collected by the experimenters, handed to B and stage four did not take place.

If at least one A decided to reduce, the entire group entered stage four. Here, all A's of the group took part in a sealed bid second price auction. The highest bidder won the right to destroy the envelope lying in front of him. Only the envelope of the winner was destroyed.[4] Note that B's payoff depends entirely on stages 1 to 3. The auction only selected the A who would be allowed to destroy the envelope, it did not affect B's payoff. The auction provides a non-arbitrary way to separate the decision to punish from the decision to punish personally. Since, in a second price auction, no participant has a reason to misrepresent his preferences, subjects are incentivized to truthfully state the value they attach to personal punishment.

[4]The minimum feasible bid was zero, the maximum feasible bid 10 and the step size 0.01. If there was a tie, the experimenters randomly chose a winner. This also applies to the special case of all three A's choosing a bid of zero.

The auction winner was informed about the second highest bid he had to pay. He could then proceed to destroy the envelope of B. The instructions did not specify any mode of destruction; however a small metal bin was present on the tables of each subject A. Additionally the experimenters placed a lighter inside the metal bin.[5] The envelopes in front of the non-winning A's were collected by the experimenters and delivered to the respective B's.

Between the test questions and the real effort task we asked some demographics from our subjects and two questions about their happiness ("how happy are you in general"/"how happy are you right now"). After stage four and before paying, we presented subjects with a second questionnaire asking their happiness again (only "how happy are you right now"), their perception of B's behavior and several attitude questions (see Appendix A.5). All subjects received a EUR 8 "show up fee" for answering the questionnaires. If a subject A had won the auction and had to pay more money than he earned in the experiment, he had to use a part of those EUR 8 to pay for his bid.

2.2 Procedures

The experiment was conducted in October 2008 in the laboratory of SFB 504 in Mannheim. We had 76 subjects in total (37% male, 63% female), who were students of various fields at University of Mannheim. The experiment consisted of four sessions; no subject participated twice. All recruitment was done via ORSEE (Greiner (2004)).

In total, the experiment lasted for about 1.5h, for which we paid an average of EUR 13.84. The full experiment was conducted via pen and paper. During the experiment, we used an experimental currency unit called "Thaler". Thaler were a printed play money handed to subjects during stages 2-7. At the end of the experiment, we exchanged all Thaler into Euro at a rate of 1:1.[6] All subjects were paid in cash and private.

3 Theoretical solutions

Before we present the results, we examine the game theoretic predictions of the classic fully selfish model and of the inequity aversion model by Fehr and Schmidt (1999) (as

[5]The subjects chose different methods to "destroy": Most ripped the envelope apart and deposited the pieces inside the metal bin. Some used their pen to cross out the envelope or wrote "destroyed" onto it. One subject used the lighter to burn the envelope inside the metal bin.

[6]The main reason for using play money was that we did not want subjects to worry about destroying legal tender.

an easy to calculate example of social preference models).

For the subgame perfect equilibrium in the classic model, rational A's could either not bid in the last stage, or, since they play a second price auction, coordinate on having just one A bid, while all others bid zero. However, not bidding is weakly dominant. In stage 3 A is indifferent between *yes* and *no*, as A's payoff is not affected by the decision. To simplify the analysis of B's decision we assume that at least two assigned A's finish stage one.[7] Given that the maximal possible punishment is EUR 8, B has a strictly dominant strategy of implementing $(2, 8)$, because her payoff is positive compared with the payoff of zero in case of $(10, 0)$. A's behavior in the first stage depends on the monetary equivalent of the effort A needs to exert to fill out the sheet. If the equivalent is below EUR 2, A strictly prefers to work.

If we assume our subjects have Fehr/Schmidt type preferences, the behavior in the last stage would be equivalent to the selfish model. In stage three, subjects would now chose *yes* after a $(2, 8)$ split by B and be indifferent after $(10, 0)$ split. For the typical values[8] Fehr and Schmidt estimated in their paper B would still decide on $(2, 8)$. During the real effort stage, more A's could now prefer not to work, if their parameter for disadvantageous inequality aversion was sufficiently high.

Behavior under the two models differs slightly in the first three stages, but in both we get the same result for the auction stage: Bidding by at most one A per group can only exist as part of unreasonable coordination equilibria and is weakly dominated by not bidding.

4 Hypotheses

We have the following three hypotheses to test.

Hypothesis 1 (Punishment) *Subjects A who receive the* $(2, 8)$ *split want the auction to happen.*

Negative reciprocity induces subjects A who experienced the bad split to reduce B's payoff. Similarly, if we assume Fehr/Schmidt type preferences subjects should also want to reduce B's payoff. Since B's payoff is automatically reduced if stage 4, that is the auction, occurs, we derive hypothesis 1.

Hypothesis 2 (Personal punishment)

(2a) *Subjects A who receive the* $(2, 8)$ *split bid positive amounts in the auction.*

[7] In the experiment all subjects A completed the real effort task (compare section 5).

[8] See table III in Fehr and Schmidt (1999).

Bid	0	> 0	Avg.(SE)	Max
Wanted punishment	55%	45%	.66 (.27)	5.5
Did not want punishment	73%	27%	.44 (.23)	3.2
Total	63%	37%	.58 (.19)	5.5

Note: Only subjects A who encountered the $(2, 8)$-split.

Table 1: Bids

(2b) *Subjects A never bid positive amounts in the auction.*

Both the classic and social preference theories predict that subjects should not care about the way in which B's payoff is reduced. On the other hand, following the reasoning put forward by Adam Smith, we could expect subjects do care about punishing personally, so we get a two-way hypothesis about behavior of the person that will punish in the auction.

Hypothesis 3 (Happiness) *Subjects A who punish personally are happier than those who do not.*

Connected to hypothesis 2b we would also expect those subjects who punish personally to have some emotional payoff from doing so that makes their monetary loss worthwhile.

5 Results

All participating subjects A earned EUR 10 in the real effort task: No subject decided not to work and all finished in time. In the allocation decision by subjects B, 16 out of 19 B implemented the $(2, 8)$ split, only 3 the $(10, 0)$ split. Since we are interested in subjects with a reason to punish, we look at the 48 subjects A who were matched with a B who chose the unfair split to test hypothesis 1. Following the $(2, 8)$ split, 58.3% of the subjects want to punish, that is, want the auction to happen. After a $(10, 0)$ split, it is demanded by only 11.1%, this is a significant difference (MW test at 0.05 level). Therefore, we affirm hypothesis 1, our subjects are seeking punishment after receiving the worse of the two allocations.

Since it is sufficient that one subject A out of the three matched to a particular B demands punishment for the auction to happen, 15 out of 16 groups where B chose (2,8) proceeded to this stage.[9] Table 1 shows the percentage of subjects A who bid a

[9] One group out of three where (10,0) was chosen also went to the auction.

positive amount in the ensuing auction - split into those who demanded punishment in the previous stage and those who did not. Recall that the auction is not payoff relevant for subject B. The payoff of B is fully determined by stages one to three. Subjects A who are either strict money maximizers or only interested in the monetary consequences of punishment for B have no incentive to bid higher than zero. In contrast to that, we find that 2/5th of our subjects bid positive amounts of money. So with respect to hypothesis 2a, we can conclude that at least a substantial minority of subjects is interested enough in punishing personally to be willing to sacrifice some of their own money to achieve this.

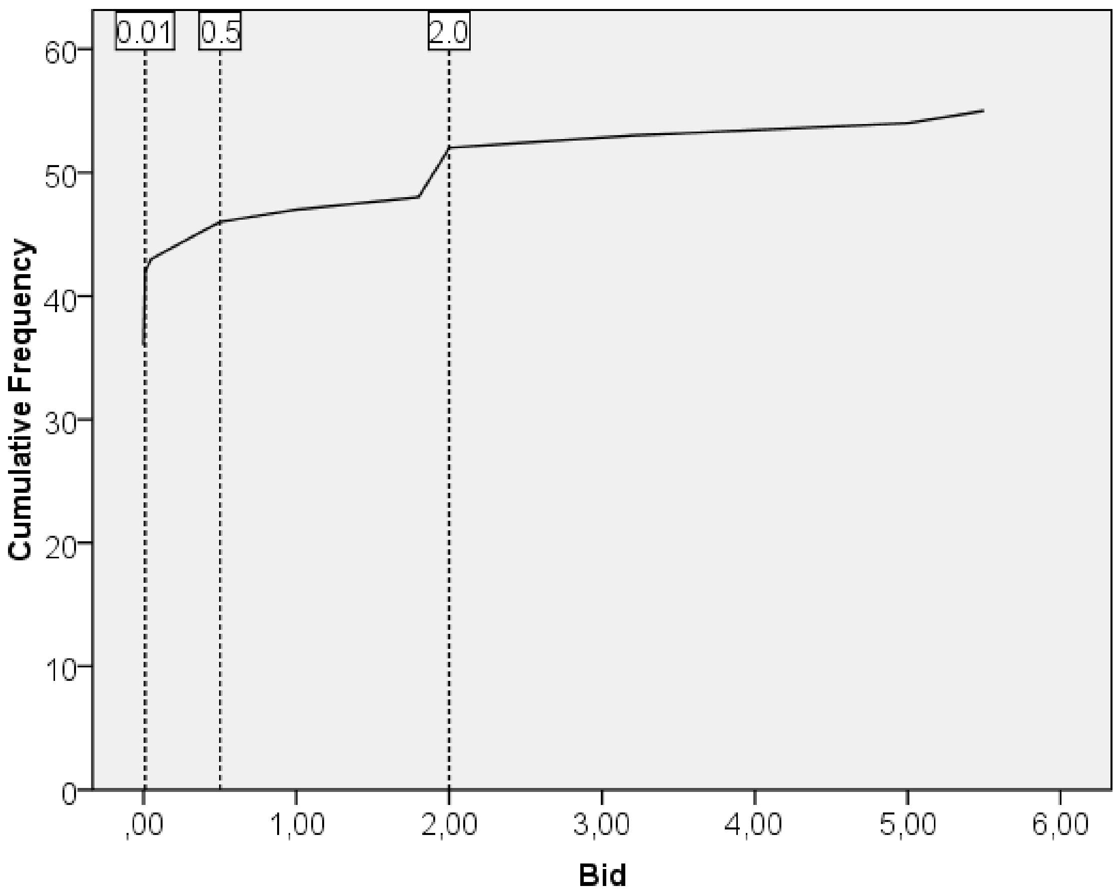


Figure 2: Cumulative distribution of bids

The cumulative distribution of bids in figure 2 shows that several subjects bid the minimum positive value of EUR 0.01, with expected jumps at focal points like EUR 0.5. The highest bid was EUR 5.5, EUR 2 was the highest price paid by any winner of the second price auction.

Surprisingly, those A who did not demand punishment in the previous stage also bid in the auction. 44.8% of the subjects who wanted the auction bid positive amounts compared to 27.3% of the subjects who did not want the auction. However, the difference is not significant according to a MW test.[10] The puzzle can be partly explained by the fact that we offered subjects only a fixed amount of punishment, so it is possible that some subjects who did not demand the auction wanted punishment in general, but did not agree with our level of punishment.

Finally, we look at the result of the physical destruction carried out by the winners of the auction. Do they enjoy the act of destroying B's money? We ask all participants for their subjective happiness on a seven point scale[11] at the start and at the end of the experiment. While the absolute level might depend on a number of causes we can not control, we can use the difference in happiness between the start and end of the experiment. Let the *happiness difference* be the amount of happiness reported at the end of our experiment minus the amount reported at the start. So subjects with a positive happiness difference felt better after our experiment than before. Not surprisingly, subjects A who encountered the allocation (10,0) felt happier compared to those who received only EUR 2. However, there is also a difference within those who were matched with a subject B who chose (2,8). Such A's who went on to win the auction have a small but positive happiness difference of on average 0.33, while it is -0.23 for those who did not win. That difference is significant on the 5% level (MW test). So despite being paid less money in the end, subjects who personally destroyed B's money leave the experiment happier that those who do not, in line with hypothesis 3.

While a "demand for personal punishment" can explain our results, there are potentially other explanations. One worry is that subjects might bid in the auction simply because they like the act of destroying the experimental money, irrelevant of the owner. To account for this, the final questionnaire included the question "Do you like destroying money?". Not one of our subjects answered with yes. Additionally, we gave subjects the opportunity to destroy some of their own remaining money during the final questionnaire. Again, none took this opportunity. Subjects might also want to use the destruction of the money as a signal of their own toughness. However, seats in the experiment were separated by blinds. Most of our subjects chose non-dramatic methods of destruction that would have been hard to notice by others. Since subjects were randomly matched, it was impossible to send a signal about ones own personality specifically to B.

[10]We also ran probit regressions for the decision to punish and bidding a positive amount, but we find our demographics and questionnaire data mostly unsignificant.

[11]See appendix A.5 for the translated questionnaires.

Yet, our findings of positive bids, of increased happiness among subjects who destroyed money and of the irrelevance of B's character traits for the amount of the bid could also be due to a desire to win auctions. To test for this possibility, we conduct a control treatment that keeps the auction format (stage 4 of the experiment described in section 2.1), but removes the punishment aspect. In both the experiment and the control treatment, subjects had to correctly answer test questions about the logic of a second prize auction before the experiment proceeded.[12]

6 Control treatment

6.1 Design

We designed the control experiment to duplicate the auction, but exclude the motivation of personal punishment. Separating the auction stage from the rest of our experiment, we had to insure that the conditions for our subjects remain comparable. We conducted the control subsequent to another, unrelated and about 1 hour long, experiment, where the subjects earned on average EUR 10.60.[13] This money was used to pay for bids in the control auction. After the end of the other experiment, we distributed the instructions for the control. Instructions and test questions were as close as possible to those in the main experiment.[14]

Subjects were placed in groups of three (corresponding to our group size of three A, who did participate in the auction), then took part in a second price auction. No subject knew the identity of the other members of the group. The highest bidder won the right to destroy an envelope lying in front of him. We handed out envelopes to all participating subjects before the bidding. In condition *full*, the envelops contained "Thaler". These "Thaler" were not used as an experimental currency, so for our subjects they were just play money or pieces of paper. The instructions stated that the auction was only for the right to destroy the envelope. However, some subjects might still have reasoned that the auction winner could keep the contents of the envelope. Therefore, in condition *empty*, we used empty envelopes instead. Feasible bids and step size were the same as in the main experiment. Only the envelope of the winner was destroyed, all others were collected by the experimenters.

The auction winner was informed about the second highest bid he had to pay. He could then destroy the envelope. At the end of the control experiment, subjects were

[12]See appendix A.5 for the translated test questions.

[13]This is close to the average earnings of EUR 11.26 that subjects of type A had accumulated in the main experiment before the auction was conducted.

[14]The control instructions correspond to stage 6-8 of the main experiment, see appendix.

Bid	0	> 0	Avg.(SD)	Max
Wanted punishment	55%	45%	.66 (.27)	5.5
Did not want punishment	73%	27%	.44 (.23)	3.2
Total Experiment	*63%*	*37%*	*.58 (.19)*	*5.5*
Control full	50%	50%	.45 (.36)	6.5
Control empty	46.7%	53.3%	.92 (.53)	6
Total control	*48.5%*	*51.5%*	*.67 (.31)*	*6.5*

Table 2: Bids control

paid privately and in cash their earnings from the prior experiment. Auction winners were paid what they earned in the prior experiment minus the second highest bid in their group.

6.2 Procedures

The experiment was conducted in June 2009 in the laboratory of SFB 504 in Mannheim. In total 33 subjects participated (18 in condition *full*, 15 in condition *empty*), mostly students of various fields at University of Mannheim. The control consisted of three sessions; no subject participated in more than one session of either main experiment or control. All recruitment was done via ORSEE (Greiner (2004)).

6.3 Results

In the control, about half of our subjects bid positive amounts. While the average bid is higher in the treatment with empty envelopes, the difference between the two control treatments is not significant (M-W test, p=0.428 two-sided). Therefore, we pool the data. Figure 3 compares the distribution of bids the main experiment and control. The higher frequency of positive bids in the control is mainly due to more subjects bidding very small positive amounts. When we compare the average bid over all subjects who take part in the auction, we find no significant difference between the pooled control and the main treatment (M-W test, p=0.327 two-sided).

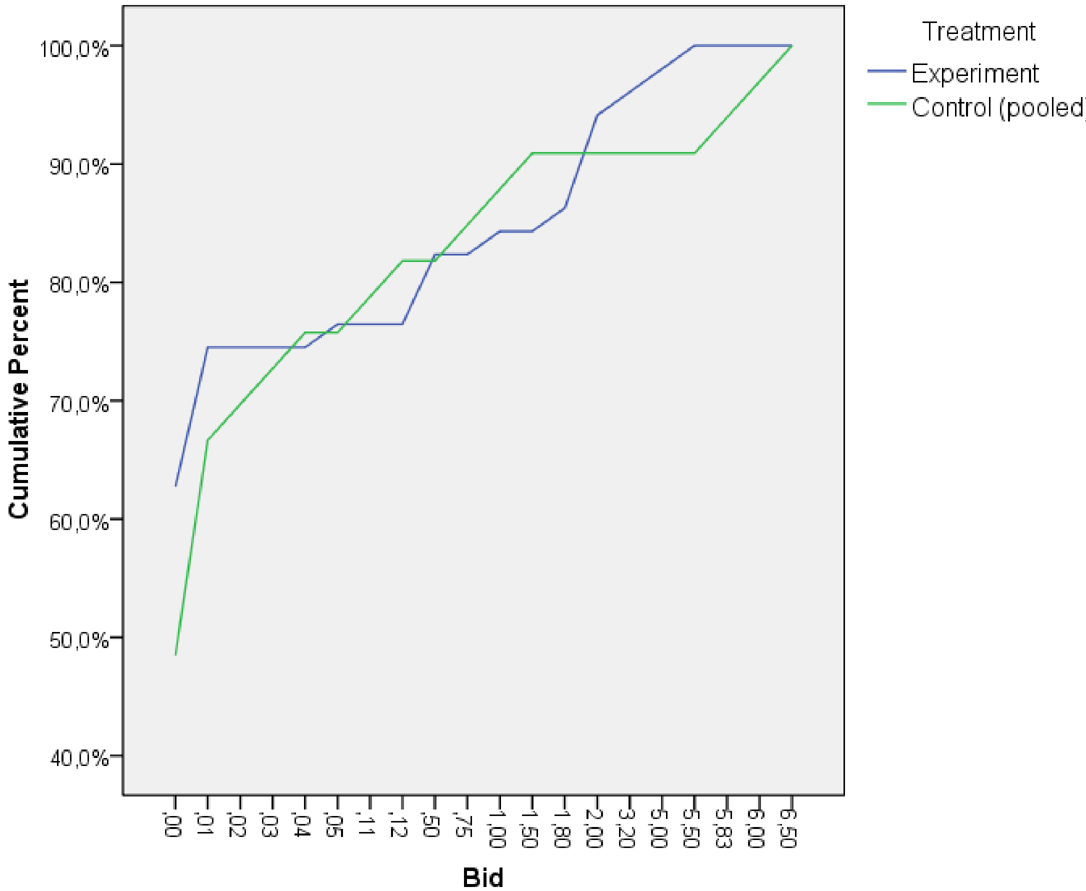


Figure 3: Cumulative distribution of bids by treatment

7 Discussion

In an experiment designed to separate the decision to punish personally from the more general decision to punish, we find that many subjects bid positive amounts in a second price auction that auctions off the right to punish personally. Some of these subjects bid substantial amounts. Consistent with positive bids, subjects who win that right report becoming happier during the experiment compared to those who do not win. While subjects punish more often based on their perception of the matched other player, the decision to bid in the auction seems to be determined only by personal character traits, not the perception of the other.

The positive bids are not due to a desire to destroy money in general. Neither can they easily be explained as being signals and we try to reduce subject's confusion as much as possible and test their understanding of a second price auction. However, our control treatment points out a second possible explanation for subjects bidding money: a "desire to win" the auction. There are some experimental papers that have looked at the issue whether winning an auction has a value in itself (apart from the value of the auctioned object), with divergent results. Holt and Sherman (1994) let subjects play auctions against computerized opponents in treatments which facilitate a "winner's curse" an opposing "loser's curse" and a balanced treatment, which they use to identify a desire to win (since in this treatment, no informational bias should occur). They do not find such behavior among their subjects. On the other hand, Ku et al. (2005) use data from real life and Internet auctions and a laboratory experiment with sequential bidding auctions to show that bidding behavior is consistent with models of escalation or competitive arousal, which could explain a desire to win. Closest to our control treatment is a study by van den Bos et al. (2008). In this experiment, subjects bid in a sealed bid first price auction. In one treatment, the opponents are other human subjects (similar to our control treatment), while in two other treatments, subjects bid against computerized agents. Furthermore, all subjects are taught to calculate the (risk-neutral) Nash-equilibrium strategy, to rule out a winner's curse effect stemming from limited cognitive ability. They find that subjects playing against humans overbid significantly more often than those playing against computers. There is also evidence from a fMRI experiment by Delgado et al. (2008) who compare subjects' reactions to losing a lottery versus losing an auction to conclude that "The fear of losing the social competition inherent in an auction may lead people to pay too high a price for the good for sale". It is possible that, in a similar vein, our subjects did not want to "lose" the auction and therefore bid positive amounts.

The concept of personal punishment has intuitive appeal. However, given the results of our control treatment, we can not conclude that the positive price our subjects are willing to pay is due to this motivation. As a final caveat, personal punishment, as Adam Smith describes it, is punishment for a grave offense. For obvious reasons, laboratory experiments can only implement minor offenses, which need not necessarily trigger the same kind of demand for personal punishment.

References

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100:464–477.

Bolton, G. E. and Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90(1):166–193.

Casari, M. and Luini, L. (2009). Cooperation under alternative punishment institutions: An experiment. *Journal of Economic Behavior & Organization*, 71(2):273 – 282.

De Quervain, D., Fischbacher, U., Treyer, V., Schnellhammer, M., Schnyder, U., Buck, A., and Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, 305:1254–1258.

Delgado, M. R., Schotter, A., Ozbay, E. Y., and Phelps, E. A. (2008). Understanding overbidding: Using the neural circuitry of reward to design economic auctions. *Science*, 321(5897):1849 – 1852.

Dufwenberg, M. and Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268 – 298.

Falk, A. and Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2):293 – 315.

Falkinger, J., Fehr, E., Gächter, S., and Winter-Ebmer, R. (2000). A simple mechanism for the efficient provision of public goods: Experimental evidence. *The American Economic Review*, 90(1):247–264.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90:980–994.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.

Greiner, B. (2004). An online recruitment system for economic experiments. In Kremer, K. and Macho, V., editors, *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63*, pages 79–93. Elsevier, Göttingen: Gesellschaft für Wissenschaftliche Datenverarbeitung.

Holt, C. A. and Sherman, R. (1994). The loser's curse. *American Economic Review*, 84(3):642 – 652.

Ku, G., Malhotra, D., and Murnighan, J. K. (2005). Towards a competitive arousal model of decision-making: A study of auction fever in live and internet auctions. *Organizational Behavior and Human Decision Processes*, 96(2):89 – 103.

Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1:593–622.

Masclet, D., Noussair, C., Tucker, S., and Villeval, M.-C. (2003). Monetary and non-monetary punishment in the voluntary contributions mechanism. *American Economic Review*, 93(1):366–380.

Nikiforakis, N. and Normann, H.-T. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics*, 11(4):358 – 369.

Ostrom, E., Walker, J., and Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *The American Political Science Review*, 86(2):404–417.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):1281 – 1302.

Smith, A. (1759). *The Theory of Moral Sentiments*. Edinburgh.

van den Bos, W., Li, J., Lau, T., Maskin, E., Cohen, J. D., Montague, P. R., and McClure, S. M. (2008). The value of victory: Social origins of the winner's curse in common value auctions. *Judgment and Decision Making*, 3(7):483 – 492.

Appendix

A Instructions Experiment

A.1 Instructions

Welcome to our experiment! Please read these instructions carefully. From now on, do not talk to your neighbors. Please turn off your mobile phone and leave it off till the end of the experiment. If you have a question, raise your hand. We will then come to you.

For the experiment, each participant is assigned one of two roles: A or B. You are in the **role of** $A[B]$[15] for the entire experiment. Three A are always matched with one B. No participant will get to know the identity of the other participants. In the experiment, we use the experimental currency unit Thaler. At the end of the experiment, the paper Thaler will be exchanged into Euro with a rate of 1 Thaler = EUR 1. Each participant will be paid private and in cash. Your payout depends on your decisions during the experiment and on the decisions of the other participants you are matched with.

Procedure of the experiment

Step 1: Questionnaire Please answer the questionnaire we hand out. You will get 8 Thaler for doing so.

Step 2: Graph paper Each participant A receives one page of graph paper and a pen. It is his task to fill this page with "+" and "o" signs according to the template. For this, he has a maximum of 25 minutes. If he fills the entire page, he produces 10 Thaler. If not, he produces 0 Thaler and does not take part in stages 3 to 6.

Step 3: Decision of B B does now decide on one of the following distributions of the Thaler produced by A in stage 2 between himself and the A's and notes this on decision sheet B. There are two possible distributions, who then are implemented for all A:

i) 2 Thaler for A and 8 Thaler for B

ii) 10 Thaler for A and 0 Thaler for B

[15] All subjects got the same instructions, up to this sentence.

In the first case, each A receives 2 Thaler. B receives 8 Thaler for each A who still takes part in stage 3 (that means 24, 16, 8 or 0 Thaler when 3, 2, 1, 0 A's are still participating). In the second case, each A receives 10 Thaler and B receives 0 Thaler in total.

Step 4: Transfer The experimenters allocate the Thaler according to the decision of B. Each A receives:

- those Thaler, that B allocated to him.
- an envelope which contains the Thaler, produced by A, that B allocated to himself.

The envelopes must not me opened.

Step 5: Decision of A's Each A decides on the following question and notes this on decision sheet A: Should one envelope be destroyed? In case of allocation i) this will reduce the payout of B by 8 Thaler. If no A answers yes, stage 6 will be does not apply and the payout of B will not be reduced.

Step 6: Auction All three A take part in this auction, with the exception of those who dropped out in stage 2. Out of the envelopes exactly one will be destroyed, two will remain. Each A can bid for the right to destroy his own envelope with the included money which B would receive from him. Only the winner of the auction may destroy the envelope. B will not receive any Thaler out of the envelope of the winner. The auction works all follows: Each A notes the amount of Thaler which he is willing to bid on decision sheet A (minimum 0 Thaler, maximum 10 Thaler, step size 0.01 Thaler). That A who bids the highest amount of Thaler wins the auction and obtains the right to destroy his envelope. However A only pays an sum equivalent to the second highest bid. The cost will be deducted at the payout. There will always be a winner of the auction. In case of several similar highest bids, the winner will be decided randomly. Note: In this type of auction, it is optimal to bid just an amount that is equivalent to what the good on offer (here: the right to destroy the envelope) is worth to oneself.

Step 7: Result of the auction The winner of the auction can now destroy the envelope and the included Thaler in arbitrary manner. Afterwards, the envelopes of those A who did not win the auction will be handed to B. B can take the money out of these envelopes.

Step 8: Questionnaire Finally, please answer the second questionnaire.

Payment Now all Thaler are exchanged into Euro. All A's have their 8 Thaler from stage 1 plus the Thaler from stage 4. The winner additionally has to pay the second highest bid. B has 8 Thaler from stage 1 plus all Thaler from the envelopes, with the exception of the destroyed envelope.

A.2 Decision sheet B

ID:_____
Please note your decisions here, as described in the instructions.

Step 3: Decision Out of the three A's working for you _____ have completed stage 2 successfully and produced 10 Thaler.
Please tick a box to mark the allocation you have decided on:
o 2 Thaler for A and 8 Thaler for B or o 10 Thaler for A and 0 Thaler for B

A.3 Decision sheet A

ID: _____
Please note your decisions here, as described in the instructions.

Step 5: Decision Should the payout of B be reduced by 8 Thaler by destroying one envelope?
yes o no o

Step 6: Auction Bid: _____ Thaler
(minimum 0 Thaler, maximum 10 Thaler, step size 0.01 Thaler)

A.4 Test Questions

Question 1: What payment will you receive at the end of the experiment, if you are A and you do not manage to fill out the complete graph paper.

Question 2: As A, you are bidding 2 Thaler in the auction. A second A bids 0 Thaler and the third A bids 5 Thaler.

a) Who wins the auction and may destroy the white envelope?

b) How much does the winner have to pay?

Question 3: Assuming all A's were successful in stage 1 and B did decide on the allocation "2 Thaler for each A and 8 Thaler for B". Look at stage 5 and 6. What is the only case in which the payout of B is not reduced by 8 Thaler?

Question 4: You are B. All A did fill out the complete paper in stage 2 and you did decide on the allocation "2 Thaler for each A and 8 Thaler for B". The A's decided they want the auction. In the auction, the A's are bidding exactly as in question 3. What is your payout at the end of the experiment?

Question 5: You are B. 2 out of 3 A's did fill out the complete paper in stage 2 and you did decide on the allocation "10 Thaler for A and 0 Thaler for B". All A's decided against the auction.

a) What payout will you receive at the end of the experiment?

b) What is the payout those A's who completed the paper will receive?

c) What is the payout of the A who did not complete the entire paper?

A.5 Questionnaire 1

ID: _____
How happy are you in general?
(very unhappy) o o o o o o o (very happy)
How happy are you at the moment?
(very unhappy) o o o o o o o (very happy)
How old are you?
What is you gender?
o M o F
Are you a student?
o yes o no
If yes: What is you major?

A.6 Questionnaire 2

ID: _____

How happy are you at the moment?

(very unhappy)	o	o	o	o	o	o	o	(very happy)

How did you perceive B's behavior in stage 3?

(not fair)	o	o	o	o	o	o	o	(fair)
(not nice)	o	o	o	o	o	o	o	(nice)
(not comprehensible)	o	o	o	o	o	o	o	(comprehensible)
(not rational)	o	o	o	o	o	o	o	(rational)
(not egoistic)	o	o	o	o	o	o	o	(egoistic)

In general, do you like destroying money?
o true o not true

I am always fair to others, even if I am at a disadvantage because of it.
o true o not true

I think fairness is an exceptionally important characteristic of humans.
o true o not true

I dislike taking responsibility.
o true o not true

I rarely hit back, even if someone else hits me first.
o true o not true

If someone hits me first, I'll show him.
o true o not true

If I am angry I occasionally bang doors shut.
o true o not true

If someone angers me, I tend to tell him what I think about him.
o true o not true

Even if I don't show it, I am sometimes consumed with envy.
o true o not true

If someone does not treat me right, I do not let it get at me.
o true o not true

Before we pay out the money, you have the possibility to destroy an arbitrary amount of your own Thaler lying in front of you. Do you want to destroy Thaler?
o Yes, _____ Thaler o No, I don't want to destroy my own Thaler.

B Instructions Control

B.1 Instructions

In this experiment, you are, together with 2 other participants, in a group of 3 people. No participant will get to know the identity of the other participants.

In front of every participant is an envelope. In this experiment, the right to destroy this envelope is auctioned off.

Auction All three participants take part in this auction. Exactly one envelope will be destroyed, two will remain. Each participant can bid for the right to destroy his own envelope. Only the winner of the auction may destroy the envelope. The auction works all follows: Each participant notes the amount of Euro which he is willing to bid on decision sheet (minimum 0 Euro, maximum 10 Euro, step size 0.01 Euro). That participant who bids the most wins the auction and obtains the right to destroy his envelope. However he only pays a sum equivalent to the second highest bid. The cost will be deducted at the payout. There will always be a winner of the auction. In case of several similar highest bids, the winner will be decided randomly. Note: In this type of auction, it is optimal to bid just an amount that is equivalent to what the good on offer (here: the right to destroy the envelope) is worth to oneself.

Result of the auction The winner of the auction can now destroy the envelope in arbitrary manner. Afterwards, the envelopes of those participants who did not win the auction will be collected by the experimenters.

Payment The winner has to pay the second highest bid. All other participants pay nothing.

B.2 Test Questions

Question 1: You are bidding 2 Euro in the auction. A second participant bids 0 Euro and the third bids 5 Euro.

a) Who wins the auction and may destroy the envelope?

b) How much does the winner have to pay?

Question 2: Is it possible in any group that in no participant destroys his envelope?

Question 3:

a) Assume you bid 0 Euro in the auction. What payment will you receive for this part of the experiment?

b) Assume you bid 0 Euro in the auction and this is the highest bid, the second highest being 1 Euro. What payment will you receive for this part of the experiment?

III. Punishment with Uncertain Outcomes in the Prisoner's Dilemma*

Abstract

This paper experimentally investigates whether risk-averse individuals punish less if the outcome of punishment is uncertain than when it is certain. We compare subjects' behavior in two treatments: Certain Punishment in which the prisoner's dilemma game is followed by a punishment stage allowing subjects to decrease the other player's payoff by 2 Euros; and Uncertain Punishment in which subjects could decrease the other player's payoff with a 50% probability by 1 Euro and with a 50% probability by 3 Euros. We observe only several instances of punishment in our setup. Consequently, we find that in both cases risk-averse subjects are equally likely to cooperate in the prisoner's dilemma and equally likely to punish in the punishment stage.

* Paper co-authored by Maroš Servátka

1. Introduction

Imagine two researchers working on a joint project. Both can work hard or free-ride on the work of the other researcher. For a certain set of outcomes this situation resembles a prisoner's dilemma. The researcher's action will eventually be revealed to his co-author, who will then have an opportunity to punish him for slacking off. What could a punishment look like? For example, it could take the form of sharing the bad experience with other colleagues in the profession. However, it is unclear what effect the punishment will have on the free-riding researcher. On one hand, it might affect a tenure decision or hiring in a close race for a job or perhaps discourage other colleagues to work with the person in the future. On the other hand, the punishment might be meaningless if other factors already determined the outcome or when other colleagues disregard the information about the free-riding researcher's input into the project.

Because the co-author who punishes does not necessarily know the free-riding researcher's situation, she cannot fully assess the impact of the punishment. Thus, the decision to punish could be viewed as having an uncertain outcome. In fact, there are many real life situations where the punishment might not "get through" at all and thus be insignificant to the recipient.[1] It is thus natural to ask whether uncertainty associated with the punishment outcome is an important determinant of the punishment decision and whether the punisher's risk attitude sheds any light on her behavior. Understanding the role of risk attitude might turn out to be socially beneficial as it is likely to affect individuals' willingness to cooperate (Becker, 1968) and potentially also punish.

Distributional models of other-regarding preferences (e.g., Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002) that are often used to explain the punishing behavior noted elsewhere in the literature do not consider uncertainty of outcomes. On their own such models cannot provide guidance on whether we should observe more or less punishment in situations when its outcome is uncertain. Moreover, in the above example the uncertainty pertains to the free-riding researcher but does not directly affect the (monetary) payoff of the decision maker. Thus, none of existing expected utility theories can be directly applied without making an additional assumption on how uncertainty affects preferences over the payoffs of the other person.

In this paper we introduce the assumption that a decision maker's risk attitude also determines preferences over expected payoffs of other people in the same manner as it determines preferences over her own expected payoff. Based on this assumption, we

[1] Sometimes punishment might also target a wrong person.

form a conjecture that risk-averse subjects punish less if the outcome of punishment is uncertain and test it experimentally.

In the experiment subjects have an opportunity to punish their counterpart, but the outcome of punishment depends on the realization of a random variable. The subjects' decisions are compared to those in another treatment in which the outcome of punishment is certain. We have embedded our explorations in a one-shot prisoner's dilemma game in which the players decide whether to defect and maximize their own monetary payoffs or to cooperate and maximize the joint surplus as we feel that the simplicity of the environment is a virtue in exploratory projects such as this one.[2] The primary reason for the one-shot horizon is to archive clear independence of subjects' decisions.

Social dilemma situations have been extensively studied in the economics literature for a long time (see Roth, 1988 for an overview). Fehr and Gächter (2000a) show that incentives to free-ride in a voluntary contribution mechanism (VCM) experiment can be counteracted by introducing a second stage which allows for punishment. Despite the punishment being costly, many subjects use that opportunity to deter defection. Initially, the effect on cooperation is small, but the contributions to a public project increase over time in a repeated game. A considerable amount of literature follows this paper and extends the result to non-pecuniary sanctions (Masclet *et al*., 2003; Noussair and Tucker, 2005) and explores the effectiveness of punishment (Nikiforakis and Normann, 2008) as well as the price of punishment (Anderson and Putterman, 2006; Carpenter (2007).

In these experiments (and all other experiments on social dilemmas which we are aware of) the outcome of the punishment is certain.[3] The literature on contract

[2] Despite a large body of experimental literature on cooperation in the prisoner's dilemma game we are not aware of any studies exploring the effect of punishment other than increasing cooperation if the interaction is repeated. In Duersch and Servátka (2010) we study whether a possibility of punishment increases cooperation (as observed in VCM experiments) by comparing the behavior of subjects in a prisoner's dilemma game without punishment with behavior in the Certain Punishment treatment presented in the current paper. We find that the mere availability of punishment does not increase cooperation as the subjects correctly anticipate that they will not be punished for defection.

[3] The strategic uncertainty on the side of the recipient can stem from subjects having heterogeneous punishment attitudes, such that some will punish in a situation where others will not. We are more interested in evaluating the other-regarding utility of the punishing subject, where this strategic uncertainty is less pronounced. The experiments which allow for counter-punishment (e.g., Denant-Boemont *et al*., 2007; Nikiforakis, 2008; Engelmann and Nikiforakis, 2008) are an exception, because the recipient of punishment can spend money to return the punishment. Thus, because of the strategic uncertainty stemming from unknown moves of other players, it is not obvious what the outcome of the original punishment will be.

enforcement which explores whether an agent will exert effort or shirk when faced with some probability of being monitored is of little help as the underlying game has a sequential rather than simultaneous structure and the experimental investigations (e.g., Fehr and Gächter, 2000b; Fehr and List, 2004, Dickinson and Villeval, 2008) focus on the effect of punishment on cooperation but not on the decision whether to punish or not and neither on its determinants.

Next we present the experimental setup and our results, followed by a short discussion. Instructions can be found in the appendix.

2. Experimental Setup

The experiment consists of two treatments: *Certain Punishment (cp)* and *Uncertain Punishment (ucp)* implemented in an across subjects design. In both treatments a prisoner's dilemma game is followed by a punishment stage. The prisoner's dilemma game payoffs (presented in Table 1) are denoted in Euros. The row player chooses Top (cooperation) or Bottom (defection), while the column player chooses Left (cooperation) or Right (defection). All information was common knowledge.

Table 1: Prisoner's Dilemma Payoffs

	Left	**Right**
Top	5,5	0,8
Bottom	8,0	2,2

In the punishment stage of the Certain Punishment treatment subjects could decrease their counterpart's payoff by 2 Euros with certainty at the cost of 1 Euro to themselves. In the Uncertain Punishment treatment subjects could decrease the other player's payoff with 50% probability by 1 Euro and with 50% probability by 3 Euros at the cost of 1 Euro to themselves.[4] Thus, while the expected punishment was the same in both cases (2 Euros), its outcome depended on the state of the world in the Uncertain

[4] Subjects who earned 0 in the prisoner's dilemma were able to inflict punishment using their show up fee.

Punishment treatment. The subjects were instructed that a coin would be flipped in front of them to determine the punishment outcome. If the coin toss lands on heads, the other player's payoff is decreased by 1 Euro. If the coins toss lands on tails, the other player's payoff is decreased by 3 Euros. If the subject decided to punish, the costs of punishment were incurred irrespective of the outcome of the probability draw.[5]

In our experiment the subjects played the game only once as a repeated environment would introduce a confounding factor of being punished by someone else in previous rounds. It should be noted that most previous experimental designs on punishment allow for repeated interaction. The decision to punish could be explained as an attempt to induce cooperation in the future. However, Fehr and Gächter (2000a) and other studies provide evidence that subjects punish in the last round and that they punish in a stranger matching when there is no chance of encountering the same subject(s) more than once. Thus, we anticipated observing a relatively large proportion of punishing subjects even in the one-shot scenario.

The expected effectiveness of punishment (2:1) in our experiment is lower than usually observed in the literature and the expected costs (1 Euro) are relatively high. These two design parameters were driven primarily by the consideration of not allowing subjects to make losses as this was the policy of the laboratory where the experiment was run. An alternative way of ensuring that subjects do not make losses would have been to significantly increase the show up fees. However, this could cause the subjects to perceive the game payoffs as relatively small compared to the constant fee. In order to get an estimate of whether the cost of punishment or the fear of "over-punishment" deterred some of our subjects from punishing, we included a couple of questions pertaining to the demand for punishment in a non-paid questionnaire administered at the end of the experiment.

To measure risk attitudes of subjects we used the method developed by Holt and Laury (2002). That is, subjects were repeatedly offered a choice between two lotteries, one involving higher risk than the other. From subjects' choices between ten lottery pairs it is possible to calculate their individual risk aversion parameter. The instructions are provided in the appendix.

[5] We decided to design the uncertain outcome in such a way that the punished person always learns that he or she is being punished. It is possible that subjects' behavior would differ if there was a chance that the punished person does not learn about the punishment. However, this was not the focus of our study and we leave it for future explorations.

2.1 Procedures

The experiment was conducted at the SonderForschungsBereich 504 laboratory at the University of Mannheim in May and June of 2009. It was run manually under a single blind social distance protocol. The experiment consisted of 10 sessions that lasted on average 50 minutes including the payment of subjects. A total of 140 students of various majors (about half either economics or business), recruited from the laboratory subject pool, participated in Certain Punishment and Uncertain Punishment treatments. Most of the students had previously participated in economics experiments. Each subject only participated in a single session of the study. Subjects earned on average 10.57 Euro, including a 5 Euro show up fee and the payout from the Holt and Laury procedure.

The sequence of events in a session was the following. (i) Upon entering the laboratory subjects drew a ball from an urn. The number that was indicated on the ball assigned their seat for the experiment and thus determined the matching which was done according to a pre-assigned matching protocol. (ii) The neutrally framed instructions (in German) for the prisoner's dilemma and the punishment stages were handed out. All sheets indicated subjects' ID numbers. (iii) The subjects read the instructions and afterwards were encouraged to ask questions. All questions were asked and answered individually. (iv) The subjects simultaneously made their decisions for the prisoner's dilemma game. (v) The experimenters collected the decisions forms, transferred the decision information to their anonymous counterparts' decision forms and returned them to subjects. This prevented the exchange of superfluous information and aided in maintaining the anonymity of individual decisions. (vi) After learning the decision of the paired player the subjects made their decisions regarding punishment on a second decision form. (vii) The experimenters collected the decision sheets for prisoner's dilemma and punishment stages.

(viii) Then the instructions and decision forms for the risk attitude elicitation task were handed out, filled out by subjects, and collected by the experimenters, one at a time. Subjects were informed beforehand that there would be an additional individual task after the prisoner's dilemma game with punishment, but not about the nature of this task. (ix) At the end of the session subjects filled out a questionnaire asking for their demand for punishment and demographics.[6] (x) Afterwards all of the subjects were

[6] Because the decision tasks were relatively simple we opted not to include test questions or examples in the instructions in order not to bias the subjects. Answering the questionnaire was not required for payment.

paid privately in cash. Each subject received the following amount: the show up fee of 5 Euro as an endowment plus the earnings in the prisoner's dilemma minus the punishment minus the punishment costs plus a payment for one randomly chosen lottery from the risk attitude elicitation task. All uncertainties and lotteries were resolved by flipping a coin/rolling a 10-sided die.[7]

2.2 Predictions and Hypotheses

In the classical solution for self-regarding preferences no punishment will ever be observed in our setup because it is costly and players will always choose to defect.[8] Via models of other regarding preferences, e.g. inequity aversion, it is possible to explain punishment. However, popular other regarding preferences models are silent on the issue of risk. How risk is to be incorporated into the utility function is not explained. Therefore, these models do not offer a clear prediction whether we should observe more punishment in our Uncertain Punishment treatment or not.

It is well known that for risk averse subjects, lotteries over outcomes are worth less than their expected value:

$$U(0.5x^{low} + 0.5x^{high}) > 0.5U(x^{low}) + 0.5U(x^{high})$$

Our research question relates to subjects' risk attitudes as predictors of their punishing behavior in the face of uncertain outcomes. Model our subjects' utility to depends on their own payoff and the punishment inflicted on the paired player, U(x,p). If subjects value punishment, and if they are risk averse, we might expect them to be risk averse with respect to the strength of punishment as well. So a comparable inequality to the one above might be:

$$U(x, 0.5p^{low} + 0.5p^{high}) > 0.5U(x, p^{low}) + 0.5U(x, p^{high})$$

Hence, we formulate the null as follows:

H0: Risk-averse subjects punish less in Uncertain Punishment than in Certain Punishment.

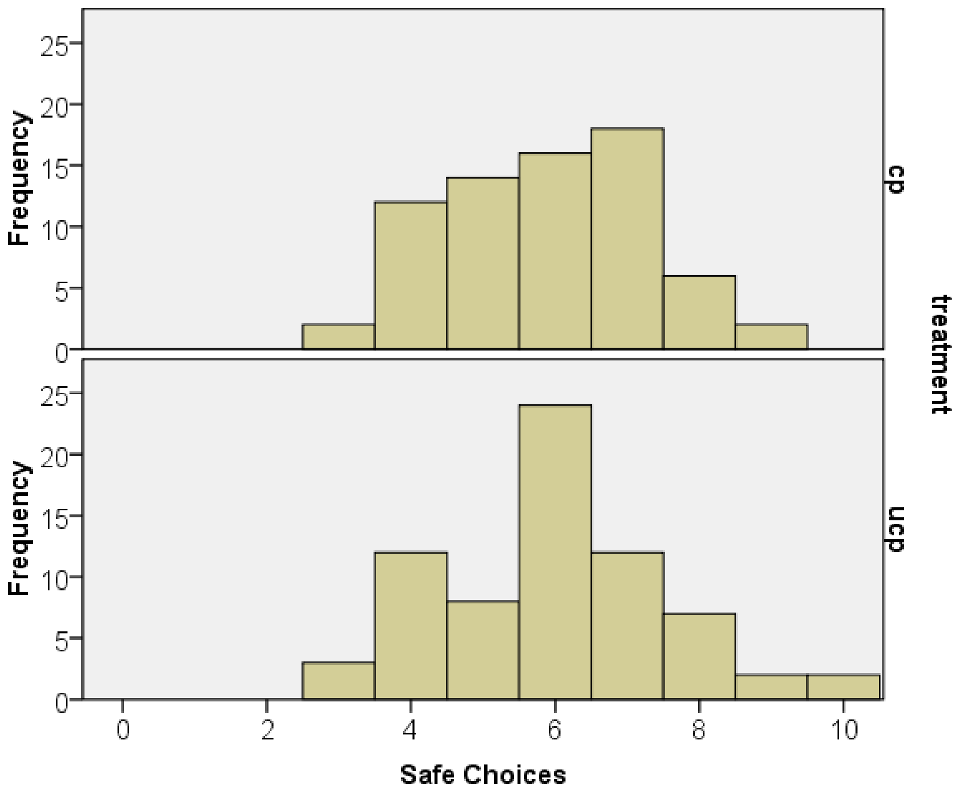
[7] The coin was flipped publicly by a randomly chosen subject. The die was rolled individually by each subject at the time of payment.

[8] Even when a self-regarding subject expects to be punished, defecting and not punishing is still a (weakly) dominant strategy.

3. Results

As our hypothesis is connected to subjects' risk aversion, we start off by describing the distribution of risk attitudes in our sample. The risk attitudes were elicited after the punishment decisions had been made but before the decision of the paired player or the punishment outcome were revealed. In the Holt and Laury method the risk attitude is determined by the number of safe choices made when choosing between a relatively safe and a risky lottery. Never choosing the safe lottery corresponds to an extremely risk-loving subject. On the other hand, the higher the number of safe choices, the more risk-averse the subject is.

Figure 1: Distribution of Risk Attitudes



The distribution of safe choices is shown in Figure 1. Risk neutrality corresponds to choosing the safe lottery exactly four times.[9] A Kolmogorov-Smirnov test on the cut-

[9] Definite statements about the risk attitude are only possible if the choices are monotonically ordered, that is when there is one lottery such that the subject always chooses the relatively safe lottery for lower

off value in the Holt and Laury task shows that there is no significant difference in the distribution of risk attitudes between the two treatments. In line with other experiments (e.g., Harrison, 1986; Cox *et al*., 1988; Holt and Laury, 2002; Eckel and Grossman, 2008) most of our subjects show a considerable amount of risk aversion while only a few are risk-loving.

Despite a relatively large number of subjects in both treatments (*n* = 70) we observe few instances of punishment: there was only 1 subject who punished in Certain Punishment and 8 subjects who punished in Uncertain Punishment. It appears that given the cost and effectiveness of punishment, the subjects were unwilling to spend resources to decrease the payoff of the other person in a one-shot game where the punishment could not lead to more cooperation in the future. However, at this stage we are not able to provide a definitive answer why our subjects did not punish.[10]

Table 2: Subjects' Behavior in the Prisoner's Dilemma Game and Punishment Stage across Treatments

		Prisoner's Dilemma		Punishment Stage
	# observations	**Cooperate**	**Defect**	**Punish**
Certain Punishment	70	23 32.9%	47 68.1%	1 1.4%
Uncertain Punishment	70	31 44.3%	39 55.7%	8 11.4%

Our main hypothesis asserts that risk-averse subjects punish less if the punishment outcome is uncertain. At the same time, the punishment decision likely depends on the history of play: When making their choice, subjects who cooperated might decide differently from those who defected and probably behave differently towards cooperators than towards defectors (data presented in Appendix 1).

ranked lottery pairs and the more risky lottery for higher ranked lottery pairs; 93.6% of our subjects display such monotonic choice behavior. In the analysis we use the data on all subjects but control for those whose choices were non-monotonic ("inconsistent") and also for those who chose the dominated safe option in the last row of the risk attitude elicitation task ("confused").

[10] It is possible that the observed behavior is driven by the large proportion of business and economics majors in our sample. Marwell and Ames (1981), Carter and Irons (1991), Frank *et al.* (1993), and Rubinstein (2006) show that economics students behave in accordance with the predictions of neoclassical theory. However, adding a dummy for business/economics students does not change our main result.

To address this issue we include a dummy variable for a subject's own behavior in the prisoner's dilemma (*OwnPD*) and for the behavior of his or her counterpart (*OtherPD*) in the probit model exploring the determinants of punishment. In the regression reported in Table 3 *Uncertain Punishment* represents a treatment dummy variable; *Risk aversion* a dummy for risk-averse subjects; *Confused* a dummy which takes on the value of 1 if the subject chose the lower paying lottery in the last row of the risk attitude elicitation task;[11] and *Age* and *Male* are the subject's age and gender as reported in the post-experimental questionnaire.

While we observe that both one's own and the paired person's behavior in the prisoner's dilemma game are important factors of the punishment decision, risk aversion, age, and gender are insignificant. If we exclude socio-demographic variables the statistical significance of *OwnPD* and *OtherPD* decreases (presented on the right hand side of Table 3). This result is robust to representing the risk attitude by the number of safe choices instead of the risk aversion dummy, to excluding inconsistent subjects, and also to including an interaction variable *UCP x RA*. Hence, we reject hypothesis H0.[12]

Table 3: Probit Regression Estimates for the Punishment Behavior

Punish	**Coefficient**	**St. Error**	***Z***	***p* > |*z*|**	**Coefficient**	**St. Error**	***Z***	***p* > |*z*|**
Own PD	0.87	0.42	2.07	0.038	0.66	0.38	1.74	0.083
Other PD	-1.11	0.55	-2.00	0.045	-0.65	0.42	-1.57	0.117
Risk Aversion	-0.49	0.45	-1.11	0.267	-0.30	0.42	-0.72	0.470
Confused	1.91	0.94	2.03	0.043				
Age	-0.03	0.06	-0.05	0.961				
Male	-0.19	0.41	-0.46	0.647				

[11] In total, there are 9 inconsistent subjects (5 in Certain Punishment and 4 in Uncertain Punishment) and 3 confused ones (all 3 in Uncertain Punishment). However, none of the inconsistent subjects punished.
[12] Because of the low number of punishment instances we also use the Fisher's exact test to verify that the proportion of cooperators who punished defectors in Certain Punishment is greater than in Uncertain Punishment. The null is rejected at p = 0.360.

Uncertain Punishment	0.97	0.51	1.88	0.060	1.03	0.48	2.17	0.030
Constant	-2.02	1.49	-1.35	0.176	-2.12	0.58	-3.68	0.000

Number of observations = 140

Recall that in order to keep our design simple we chose to restrict the punishment to only one (expected) option. However, it is possible that some subjects would like to punish more or less and thus the size of desired punishment might vary across treatments. In order to get at least partially at this issue, in the post-experimental questionnaire we asked our subjects the following question:

If you could decide how much to destroy of the other player's payoff how much would you like to destroy?

About 20% of subjects provided a positive number as their answer. The answers do not significantly vary between treatments, suggesting that the size of the desired punishment does not interact with uncertainty of its outcome. When we treat the subjects' answers as observations, a tobit regression (presented in Appendix 2) shows a similar pattern: Risk-averse subjects did not want to destroy more or less in the Uncertain Punishment than in the Certain Punishment, thus providing further support for rejecting H0.

We also verify the effect of risk aversion on the decision to cooperate. The result supports our finding reported in Duersch and Servátka (2010) that risk aversion does not influence subjects' willingness to cooperate.

4. Discussion

This paper reports an experiment designed to study the role of risk attitude in punishing behavior when the outcome of punishment is uncertain. We assume that if the decision maker is risk-averse, her behavior will exhibit risk-aversion also over the payoffs of the person who is being punished. We observe a relatively low number of punishment instances in our setup and find no evidence that risk aversion is a factor when making a decision to punish.

There are several potential explanations and implications for our findings. We discuss three of them which are directly related to our design. The first one is that our assumption about risk aversion does not reflect reality. Based on models of conditional other-regarding preferences (e.g. Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006; Cox *et al.*, 2007; Cox *et al.*, 2008) if the decision maker decides to punish, it can be argued that the aim of such action is not to benefit the other person but to hurt him. As our design does not allow for various levels of punishment, it is possible that a decision maker takes into account the uncertainty affecting the punished person in order to hurt him more. Such approach could explain the seemingly higher number of instances of uncertain punishment: Assuming that the recipient of punishment is risk-averse, the uncertain punishment is a stronger punishment than the certain one.

The second explanation is connected with the use of Holt and Laury's measure of risk attitudes which might be not reflect the same *type of risk attitude* present in punishment decisions (Isaac and James, 2000; Dave *et al.*, 2007; Deck *et al.*, 2009). Thus, a robustness check of our findings with respect to a different risk attitude elicitation method seems warranted.

Finally, in our experiment the punished person always learns that he or she is being punished. It is possible that our results do not directly apply to environments in which there is a chance that the punished person does not find out about the punishment.

References

ANDERSON, C. M. and PUTTERMAN, L. (2006) "Do Non-Strategic Sanctions Obey the Law of Demand? The Demand for Punishment in the Voluntary Contribution Mechanism," *Games and Economic Behavior*, 54, 1-24.

BECKER, G. (1968) "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76(2), 169-217.

BOLTON, G. and OCKENFELS, A. (2000) "ERC: A Theory of Equity, Reciprocity, and Cooperation," *American Economic Review* 90, 166-193.

CARPENTER, J. (2007): "The Demand for Punishment," *Journal of Economic Behavior and Organization,* 62(4), 522-542.

CARTER, J. and IRONS, M. (1991) "Are Economists Different, and If So, Why?" *Journal of Economic Perspectives*, 5(2), 171-177.

CHARNESS, G. and RABIN, M. (2002) "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117(3), 817-869.

COX, J.C., FRIEDMAN, D., and GJERSTAD, S. (2007) "A Tractable Model of Reciprocity and Fairness," *Games and Economic Behavior*, 59, 17-45.

COX, J.C., FRIEDMAN, D. and SADIRAJ, V. (2008) "Revealed Altruism," *Econometrica*, 76, 31-69.

COX, J.C., SMITH, V., and WALKER, J. (1988) "Theory and Individual Behavior of First Price Auctions," *Journal of Risk and Uncertainty*, 61-99.

DAVE, C., ECKEL, C., JOHNSON, C., and ROJAS, C. (2007) "On the Heterogeneity, Stability, and Validity of Risk Preference Measures," University of Texas at Dallas working paper.

DECK, C., LEE, J., REYES, J., and ROSEN, C. (2009) "Measuring Risk Attitudes Controlling for Personality Traits" University of Arkansas working paper.

DENANT-BOEMONT, L., MASCLET, D., and NOUSSAIR, C. (2007) "Punishment, Counterpunishment and Sanction Enforcement in a Social Dilemma Experiment," *Economic Theory*, 33, 145–167.

DICKINSON, D. and VILLEVAL, M.-C. (2008) "Does monitoring decrease work effort? The complementarity between agency and crowding-out theories," *Games and Economic Behavior*, 63, 56-76.

DUERSCH, P. and SERVÁTKA, M. (2010) "(No) Punishment in One-Shot Prisoner's Dilemma," University of Canterbury working paper.

DUFWENBERG, M. and KIRCHSTEIGER, G. (2004) "A Theory of Sequential Reciprocity," *Games and Economic Behavior*, 47, 268-298.

ENGELMANN, D. and NIKIFORAKIS, N. (2008) "Feuds in the Laboratory? A Social Dilemma Experiment," Dept. Economics, University of Melbourne, Research Paper 1058.

ECKEL, C. C., and GROSSMAN, P.J. (2008) "Forecasting Risk Attitudes: An Experimental Study Using Actual and Forecast Gamble Choices," *Journal of Economic Behavior and Organization,* 68(1), 1-17.

FALK, A. and FISCHBACHER, U. (2006) "A Theory of Reciprocity," *Games and Economic Behavior*, 54, 293-315.

FEHR, E. and GÄCHTER, S. (2000a) "Cooperation and Punishment in Public Goods Experiments," *The American Economic Review*, 90, 980-994.

FEHR, E. and GÄCHTER, S. (2000b) "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives*, 14, 159-181.

FEHR, E. and LIST, J. A. (2004) "The Hidden Costs and Returns of Incentives – Trust and Trustworthiness among CEOs," *Journal of the European Economic Association*, 2(5), 743-771.

FEHR, E. and SCHMIDT, K. (1999) "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics,* 114, 817-868.

FRANK, R., GILOVICH, T. and REGAN, D. (1993) "Does Studying Economics Inhibit Cooperation?" *Journal of Economic Perspectives*, 7(2), 159-171.

HARRISON, G. (1986) "An Experimental Test for Risk Aversion," *Economics Letters*, 21: 7-11.

HOLT, C. A. and LAURY, S. K. (2002) "Risk Aversion and Incentive Effects," *The American Economic Review*, 92, 1644-1655.

ISAAC, R. M. and JAMES, D. (2000) "Just Who Are You Calling Risk Averse," *Journal of Risk and Uncertainty*, 20(2), 177-87.

MARWELL, G. and AMES, R. (1981) "Economists Free Ride, Does Anyone Else? Experiments on the Provision of Public Goods," *Journal of Public Economics*, 15(3), 295-310.

MASCLET, D., NOUSSAIR, C, TUCKER, S., and VILLEVAL, M. C., (2003): "Monetary and Non-monetary Punishment in the Voluntary Contributions Mechanism," *American Economic Review*, 93, 366-380.

NIKIFORAKIS, N. (2008) "Punishment and Counter-Punishment in Public Good Games: Can We Really Govern Ourselves?" *Journal of Public Economics,* 92, 91-112.

NIKIFORAKIS, N. and NORMAN, H.-T. (2008) "A Comparative Statics Analysis of Punishment in Public-Good Experiments," *Experimental Economics*, 11, 358-369.

NOUSSAIR, C. and TUCKER, S. (2005) "Combining Monetary and Social Sanctions to Promote Cooperation," *Economic Inquiry*, Vol. 43, 649-660.

ROTH, A. (1988) "Laboratory Experimentation in Economics: A Methodological Overview," *The Economic Journal*, 98, 974-1031.

RUBINSTEIN, A. (2006) "A Sceptic's Comment on Studying Economics," *Economic Journal*, 116, c1-c9.

Appendix 1.

Table 4: Subjects' Punishment Decisions Based on the History of Play

	Own Behavior	Other Person's Behavior	Punishment
Certain Punishment	Cooperate	Cooperate	
		Defect	1
	Defect	Cooperate	
		Defect	
Uncertain Punishment	Cooperate	Cooperate	1
		Defect	4
	Defect	Cooperate	1
		Defect	2

Appendix 2.

Table 5: Tobit Regression Estimates for Destroy

Destroy	**Coefficient**	**St. Error**	***t***	***p*** > |***t***|
Own PD	9.22	2.37	3.89	0.000
Other PD	-5.68	2.31	-2.46	0.015
Risk aversion	0.72	2.45	0.30	0.768
Confused	3.69	6.53	0.57	0.573
Age	-0.27	0.29	-0.94	0.349
Male	1.80	2.05	0.88	0.349
Uncertain Punishment	2.84	2.11	1.35	0.180
Constant	-6.01	7.59	-0.79	0.430

Number of observations = 138, left-censored at 0.

Appendix 3. UNCERTAIN PUNISHMENT INSTRUCTIONS

(Translation from German, for column players. The decision forms were printed on separate sheets. The original instructions are available from the authors upon request.)

GENERAL INSTRUCTIONS

No talking: Now that the experiment has begun, we ask that you do not talk or communicate any longer with each other. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Monetary payments: The experiment will consist of two stages and will be followed by a separate decision problem for which you will get paid as well. The amount of money you make will depend on the choices made (as described below). Each participant will receive a lump sum payment of **5 Euro.** This one-off payment can be used to pay for eventual losses. Your earnings will be paid to you in cash individually and privately at the end of the session.

Matching: During the session you will be matched with another person. However, no participant will ever know the identity of the person he or she is matched with.

Roles: In the experiment a "row" player is always randomly matched with a "column" player. You are the **column player**.

INSTRUCTIONS – STAGE 1

Your decision in Stage 1: On the DECISION FORM 1 you will see a payoff table. The row player decides between Top and Bottom rows, and the column player decides between the Left and Right columns. The intersection of the designated row and column determines which part of the payoff matrix is relevant (Top Left, Top Right, Bottom Left, Bottom Right) and thus determines the earnings for each person. In each cell, the row player's payoff is shown first and the column player's payoff is shown second. Your payoff is printed in bold.

After you have made the decision, we will collect the decision forms and inform you about the decision of your matched row player.

DECISION FORM – STAGE 1

	Left	**Right**
Top	$5, **$5**	$0, **$8**
Bottom	$8, **$0**	$2, **$2**

Please **circle** either the **Left** or the **Right** column.

INSTRUCTIONS – STAGE 2

Your Decision in Stage 2: After learning the other player's decision in Stage 1, you can **decrease** the other player's payoff with 50% probability by 1 Euro and with 50% probability by 3 Euro at the cost of 1 Euro to you. We will flip a coin in front of you to determine the outcome. If the *heads* comes up, the other player's payoff will decrease by 1 Euro. If the *tails* comes up, the other player's payoff will decrease by 3 Euro.

If you decide to **decrease** the other player's payoff, you will **circle** the words "**I want to decrease the other player's payoff."**

If you decide to **not change** the other player's payoff, you will **circle** the words "**I do not want to change the other player's payoff."**

The other player can also decrease your payoff or leave it unchanged.

DECISION FORM – STAGE 2

Do you want to decrease the other player's payoff by 1 Euro with probability 50% and by 3 Euros with probability 50% at the cost of 1 Euro to you? Please circle.

I want to decrease the other player's payoff.

OR

I do not want to change the other player's payoff.

RISK ATTITUDE ELICITATION

The next page shows ten decision questions. Each decision is a paired choice between "Option A" and "Option B."

You will make ten choices and record these in the box to the left of the option. That is, if you prefer option A to option B, you will mark an X by option A. Only one of the ten decisions will be used in the end to determine your earnings for this part of the experiment.

A ten-sided die will be used to determine payoffs; the faces are numbered from 1 to 10 (the "0" face of the die will serve as 10.) After you have made all of your choices, you will throw this die twice, once to select one of the ten decisions to be used, and a second time to determine what your payoff is for the option you chose, A or B, for the particular decision selected. Even though you will make ten decisions, only one of these will end up affecting your earnings, but you will not know in advance which decision will be used. Obviously, each decision has an equal chance of being used in the end.

Now, please look at Decision 1 at the top. Option A pays $2.00 if the throw of the ten sided die is 1, and it pays $1.60 if the throw is 2-10. Option B yields $3.85 if the throw of the die is 1, and it pays $0.10 if the throw is 2-10. The other Decisions are similar, except that as you move down the table, the chances of the higher payoff for each option increase. In fact, for Decision 10 in the bottom row, the die will not be needed since each option pays the highest payoff for sure, so your choice here is between $2.00 Euro or $3.85.

To summarize, you will make ten choices: for each decision row you will have to choose between Option A and Option B. You may choose A for some decision rows and B for other rows, and you may change your decisions and make them in any order.

When you are finished, we will collect your decision sheet. Again, two persons from the class will be randomly selected to receive the monetary payoffs. To determine the payoffs from this task you will throw the ten-sided die to select which of the ten Decisions will be used. Then you will throw the die again to determine the money earnings for the Option you chose for that Decision. If you are selected, earnings (in $) for this choice will be paid to you in cash when we finish.

So now please look at the empty boxes on the record sheet. You will have to mark an X in one and only one of the boxes in each row, depending whether you prefer option A or option B. Then the die throw will determine which of the ten decisions is going to count. We will look at the decision that you made for the one that counts, and circle it, before throwing the die again to determine your earnings for this part.

DECISION FORM

Option A		Option B
☐ 2.00€ with probability of 1/10, 1.60€ with probability of 9/10	*OR*	☐ 3.85€ with probability of 1/10, 0.10€ with probability of 9/10
☐ 2.00€with probability of 2/10, 1.60€ with probability of 8/10	*OR*	☐ 3.85€ with probability of 2/10, 0.10€ with probability of 8/10
☐ 2.00€ with probability of 3/10, 1.60€ with probability of 7/10	*OR*	☐ 3.85€ with probability of 3/10, 0.10€ with probability of 7/10
☐ 2.00€ with probability of 4/10, 1.60€ with probability of 6/10	*OR*	☐ 3.85€ with probability of 4/10, 0.10€ with probability of 6/10
☐ 2.00€ with probability of 5/10, 1.60€ with probability of 5/10	*OR*	☐ 3.85€ with probability of 5/10, 0.10€ with probability of 5/10
☐ 2.00€ with probability of 6/10, 1.60€ with probability of 4/10	*OR*	☐ 3.85€ with probability of 6/10, 0.10€ with probability of 4/10
☐ 2.00€ with probability of 7/10, 1.60€ with probability of 3/10	*OR*	☐ 3.85€ with probability of 7/10, 0.10€ with probability of 3/10
☐ 2.00€ with probability of 8/10, 1.60€ with probability of 2/10	*OR*	☐ 3.85€ with probability of 8/10, 0.10€ with probability of 2/10
☐ 2.00€ with probability of 9/10, 1.60€ with probability of 1/10	*OR*	☐ 3.85€ with probability of 9/10, 0.10€ with probability of 1/10
☐ 2.00€ with probability of 10/10, 1.60€ with probability of 0/10	*OR*	☐ 3.85€ with probability of 10/10, 0.10€ with probability of 0/10

QUESTIONNAIRE

Thank you for participating in the experiment. Finally, please answer the following questions. Your answers will have no impact on your final payoff.

1. If you could decide how much to destroy of the other player's payoff how much would you like to destroy?
2. How much of your own payoff would you be willing to pay for it?
3. How old are you?
4. What is your gender?
5. What is your major?
6. In which country were you born/raised?

IV. (No) Punishment in the One-Shot Prisoner's Dilemma*

Abstract

Does a mere availability of punishment increase cooperation in the one-shot prisoner's dilemma game? In our experiment we observe that the subjects almost never use punishment. Consistently, the data shows no increase in the cooperation rate as the subjects correctly anticipated that they would not be punished for defection. Thus, the availability of punishment is ineffective in inducing cooperation in a one-shot game. Moreover, we do not find any evidence that risk attitude is a factor when making the decision to cooperate.

* Paper co-authored by Maroš Servátka

1. Introduction

Does a mere availability of punishment increase cooperation? Ostrom *et al.* (1992) point out that allowing for sanctions imposed by peers increases cooperation in interactions where the individual interest is in conflict with the group interest and therefore, leads to socially superior outcomes. However, the punishment changes the strategic nature of the game and thus it is not obvious whether it is the altered incentives or the mere availability of punishment that shape the behavior. Moreover, many social and economic interactions are repeated in nature, confounding the effect of punishment with reputation building. To address these two issues we design an experiment that preserves defection as the dominant action and eliminates the opportunity of enforcing cooperation in future play.

Since Ostrom *et al.* (1992) a wide body of experimental literature has focused on the viability of various punishment schemes (either experimenter-imposed as in Falkinger *et al*., 2000; Dickinson, 2001; Croson *et al*., 2006 or participant-imposed as in the strand of literature that originated with Fehr and Gächter, 2000) to enhance cooperative efforts. In the participant-imposed punishment experiments the punishment usually is costly, yet many subjects use that opportunity to deter defection, as predicted by distributional models of other-regarding preferences (e.g., Fehr and Schmidt 1999; Bolton and Ockenfels 2000; Charness and Rabin 2002). The presence of a sanctioning mechanism thus enables reaping the benefits of cooperation in the long-run (Gächter, *et al.* 2008).

Most of the previous research on punishment is set in a repeated voluntary contribution mechanism (VCM) and does not provide a clear-cut answer to our research question.[1] To the best of our knowledge, the only study on punishment in a one-shot VCM setting has been done by Walker and Halloran (2004) who find that having the opportunity to punishment has no impact on the level of cooperation. However, the costly punishment is a public good itself and several studies (e.g. Fehr and Gächter, 2002) provide evidence that not all individuals punish. Therefore, it is possible that Walker and Halloran do not observe an increase in cooperation because subjects anticipate the other group members to free ride on punishment. In our experiment we eliminate this consideration by focusing on groups of only two players.

While it would be possible to construct a one-shot two-player VCM scenario with punishment, we believe that simplicity of experimental design is a virtue and therefore

[1] The primary reason for the one-shot horizon is the necessity of independence of subjects' decisions on the previous play.

we have embedded our explorations in a prisoner's dilemma game in which the players have only two actions. Each player decides whether to defect and maximize his own monetary payoff or to cooperate and maximize the joint surplus (see Roth, 1988).[2]

Finally, the (strategic) uncertainty associated with a chance of being punished is likely to be an important determinant of cooperating behavior. Becker (1968) assumes that subjects use the expected probability of being punished when evaluating their actions. Based on this motivation we also explore in our experiment whether the decision-maker's risk attitude sheds any light on individuals' willingness to cooperate.

2. The Experiment

Out experiment consists of two treatments, *Baseline* and *Punishment,* implemented in an across subjects design. In each treatment a prisoner's dilemma game is played. The game payoffs (presented in Table 1) are denoted in Euros. The row player chooses Top (cooperation) or Bottom (defection), while the column player chooses Left (cooperation) or Right (defection).

Table 1: Prisoner's Dilemma Payoffs

	Left	**Right**
Top	5,5	0,8
Bottom	8,0	2,2

In *Punishment*, after being notified of the result of the prisoner's dilemma game, subjects could decrease their counterpart's payoff by 2 Euros at the cost of 1 Euro to themselves; in Baseline there was no punishment stage.[3] All information was common knowledge.

[2] Despite the extensive literature on cooperation in the prisoner's dilemma and numerous studies on punishment in the repeated VCM, we are not aware of any studies exploring the effect of punishment in a one-shot prisoner's dilemma. Although the prisoner's dilemma incorporates motivations present also in the VCM there is no experimental evidence whether (the availability of) punishment increases cooperation. Gangadharan and Nikiforakis (2009) bridge the literature between a repeated prisoner's dilemma and VCM without punishment. They find that subjects behave more cooperatively in the prisoner's dilemma than in the VCM if they are in a group of four players but do not find a difference if the two games are played in pairs

[3] Subjects who earned 0 in the prisoner's dilemma were able to inflict punishment using their show up fee.

The demand for punishment depends on the price subjects have to pay for destroying a unit of the other person's monetary payoff. Carpenter (2002), Falk *et al.* (2005) and Nikiforakis and Normann (2008) all find that more effective punishment technologies lead to more punishment. Our effectiveness of 2:1 is slightly lower than typically used in the VCM literature, but the occurrence of punishment is still reported for even lower effectiveness than ours. This design was driven by two main considerations: (1) If the punishment is inflicted, it should be significantly damaging (costly) relative to the recipient's (punisher's) payoffs from the experiment. (2) At the same time we wanted to ensure that subjects would not make losses, as this was the policy of the laboratory where the experiment was run. An alternative way of avoiding the subjects making losses was to increase the show up fees. However, this could potentially cause the subjects to perceive the game payoffs as relatively small and alter their behavior.

The experiment was conducted at the SonderForschungsBereich 504 laboratory at the University of Mannheim in May and June of 2009. It was run manually under a single blind social distance protocol. The experiment consisted of 7 sessions (3 in Baseline and 4 in Punishment) that lasted on average 40 minutes including the payment of subjects. A total of 114 students of various majors (about half either economics or business) participated as subjects. Most of the students had previously participated in economics experiments. Subjects earned on average 10.55 Euro including a 5 Euro show up fee.

Upon entering the laboratory subjects were randomly assigned a seat and paired according to a pre-assigned matching protocol. The instructions were neutrally framed and subjects were encouraged to ask questions. The questions were asked and answered individually. After subjects had made their Baseline/Punishment treatment decisions, the Holt and Laury (2002) risk attitude elicitation task was conducted. Subjects were informed beforehand that there would be an additional individual task, but not about the nature of this task. At the end of the session subjects were asked to answer a questionnaire.[4]

All subjects were paid privately and individually. Each subject received the following amount: an endowment of 5 Euro plus the earnings in the prisoner's dilemma minus the punishment minus the punishment costs plus a payment for one randomly

[4] The decision tasks were relatively simple and in order not to bias the subjects we opted not to include test questions or examples in the instructions. Answering the questionnaire was not a requirement for payment.

chosen lottery from the risk attitude elicitation task. All lotteries were resolved by rolling a 10-sided die at the time of the payment.

3. Hypotheses

For self-regarding subjects, defecting in the prisoner's dilemma and not punishing in the second stage is a dominant strategy.

Hypothesis 1a: Cooperation (self-regarding)

Subjects will never cooperate and never punish.

The Fehr and Schmidt (1999) model of inequality aversion that has been prominently used in the punishment literature to explain the behavior of subjects, offers a different prediction: For sufficiently big values of the parameter for disadvantageous inequality aversion, α, subjects will punish. Similarly, subjects with high values of advantageous inequality aversion, β, will cooperate in the prisoner's dilemma.[5]

Hypothesis 1b: Cooperation (inequality-averse)

Strongly inequality-averse subjects will cooperate in the prisoner's dilemma and punish defectors.

How does the strategic uncertainty of the game influence risk-averse subjects? If subjects expect that some of the other subjects are strongly inequality-averse and punish defectors, while others are self-regarding or only weakly inequality-averse, they will form subjective beliefs about the probability of being punished for defection. Then, due to the uncertainty, the payoff from defection is smaller for risk-averse subjects than for risk-neutral ones. On the other hand, the payoff from cooperation is similar for both, assuming that the probability of being punished for cooperating is zero.

Hypothesis 2: Risk-aversion

In Punishment, risk-averse subjects will cooperate more than risk-neutral ones.

[5] Subjects with $\alpha > 1$ prefer to punish. Fehr and Schmidt suggest that the value is typically equal or smaller than one. However, it might be greater than one for some people. Subjects will cooperate against a cooperator, if their value of β is larger than 3/8 (ignoring the punishment). Of course, in the experiment, subjects do not know whether they play against a cooperator or a defector. A positive belief of playing against a defector raises the β needed for cooperation to be optimal.

4. Results

Table 2 presents subjects' behavior in the prisoner's dilemma game and in the punishment stage. In Baseline 43.2% of subjects chose to cooperate with their partner. The cooperation rate decreased to 32.9% with the introduction of punishment.[6] However, this decrease is not statistically significant (2-sided Fisher exact test *p*-value = 0.360). Similarly, there is no significant difference in the cooperation rate between risk averse and non-risk averse subjects (p = 0.815). We observe only one instance of punishment despite a relatively large number of subjects in the punishment treatment (n = 70).

Our subjects' behavior supports the prediction of the self-regarding preferences model with regard to punishment, as the subjects were unwilling to spend resources to decrease the payoff of the other person in a one-shot game. However, self-regarding subjects would never cooperate in the prisoner's dilemma. The observed behavior is thus consistent with the predictions of the Fehr and Schmidt model of inequality aversion given values of α lower than 1 and a range of values of β. No change in the cooperation rate in conjunction with the lack of punishing behavior suggests that subjects rationally expected that they would not be punished for defection.

Table 2: Subjects' Behavior in the Prisoner's Dilemma Game and Punishment Stage

		Prisoner's Dilemma		Punishment Stage
	# observations	**Cooperate**	**Defect**	**Punish**
Baseline	44	19 43.2%	25 56.8%	-
Punishment	70	23 32.9%	47 68.1%	1 1.4%

To assess the effect of risk aversion on cooperation, we run a probit regression and report the results in Table 3.[7] *Punishment* is a dummy variable for the respective treatment; *Risk Attitude* is the number of safe choices in the risk attitude elicitation task; *Risk Attitude*Punishment* is the interaction term between risk attitude and the

[6] It is possible that the observed behavior is driven by the large proportion of business and economics majors in our sample (53 out of 114 subjects study business, economics or a closely related field).

[7] As expected under random treatment allocation of subjects, a Kolmogorov-Smirnov test on the distribution of the cut-off value in the Holt and Laury task shows that there is no significant difference in the distribution of risk attitudes between the two treatments.

punishment treatment; *Inconsistent* is a dummy which takes on the value of 1 if the subject "jumped" back and forth between lotteries in the risk attitude elicitation task;[8] *Age* and *Male* are the subject's age and gender as reported in the post-experimental questionnaire. The estimated coefficients are presented in the first column: We find no significant effect for the interaction term and thus no support for Hypothesis 2. This is robust to excluding the demographic variables (results presented on the right hand side of Table 3) and to representing the risk attitude by a dummy for risk averse subjects instead of the number of safe choices.[9]

Table 3: Probit Regression Estimates for the Cooperation Behavior

Cooperate	**Coef.**	**St. Error**	***Z***	***p* > |*z*|**	**Coef.**	**St. Error**	***Z***	***p* > |*z*|**
Punishment	-1.33	1.01	-1.31	.189	-1.30	.99	-1.31	.189
Risk Attitude	-.11	.12	-.89	.373	-.10	.12	-.85	.393
Risk Attitude *Punishment	.18	.17	1.06	.290	.18	.17	1.08	.282
Inconsistent	.38	.54	0.71	.481				
Age	-.02	.03	-0.61	.545				
Male	-.08	.26	-0.33	.743				
Constant	.89	.99	0.90	.368	-.42	.72	.59	0.557

Number of observations = 114

5. Discussion

This paper reports an experiment designed to study the effect of punishment on cooperation in the one-shot prisoner's dilemma. We observe that our subjects almost never use punishment. Consistent with this finding, including a punishment stage had no significant effect on the cooperation rate, suggesting that subjects correctly anticipated that they would not be punished for defection. Our results thus point out that

[8] In total, there are 6 inconsistent subjects (1 in Baseline and 5 in Punishment).

[9] For the obvious reason of having no data on punishment we cannot study the effect of risk attitude on the decision to punish.

the mere availability of punishment is ineffective in inducing more cooperation. Unlike us, Sutter *et al.* (forthcoming) find that giving subjects the opportunity of punishment as an institutional choice suffices to promote cooperation in a repeated VCM even if subjects choose not to punish. Although the experimental designs are too different to draw clear-cut conclusions, Sutter *et al.*'s study once again suggests that the time horizon is crucial for cooperation.

At the first glance, the low occurrence of punishment in our setup seems to be at odds with the literature on repeated public goods VCM where punishment successfully deters free riding. As mentioned above, Walker and Halloran (2004) also find that in one-shot VCM the punishment threat does not affect cooperation rates. It appears that subjects are reluctant to waste their resources on punishment if there is no future. However, studies by Carpenter (2007) and Nikiforakis and Normann (2008) suggest that the demand for punishment and the effectiveness of punishment are important determinants of subjects' decisions and therefore call for more research seems warranted.

Finally, we do not find any evidence that risk attitude is a factor when making a decision to cooperate. This second finding might be due to the anticipated absence of punishment and/or connected with the use of Holt and Laury's measure of risk attitudes that might be not appropriate for punishment decisions such as the one presented in this paper. Studies by Isaac and James (2000), Eckel and Wilson (2004) and Dave *et al.* (2007) point out that different techniques of measuring risk attitudes yield significantly different estimates and a recent paper by Deck *et al.* (2009) finds that their subjects behave as though Holt and Laury task was an investment decision.

References

BECKER, G. (1968) "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76(2), 169-217.

BOLTON, G. and OCKENFELS, A. (2000) "ERC: A Theory of Equity, Reciprocity, and Cooperation," *American Economic Review* 90, 166-193.

CARPENTER, J. (2007): "The Demand for Punishment," *Journal of Economic Behavior and Organization,* 62(4), 522-542.

CHARNESS, G. and RABIN, M. (2002) "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117(3), 817-869.

CROSON, R., FATÁS, E. and NEUGEBAUER, T. (2007) "Excludability and Contribution: A Laboratory Study in Team Production," Wharton working paper.

DAVE, C., ECKEL, C., JOHNSON, C., and ROJAS, C. (2007) "On the Heterogeneity, Stability, and Validity of Risk Preference Measures," University of Texas at Dallas working paper.

DECK, C., LEE, J., REYES, J., and ROSEN, C. (2009) "Measuring Risk Attitudes Controlling for Personality Traits" University of Arkansas working paper.

DICKINSON, D. L. (2001) "The Carrot vs. the Stick in Work Group Motivation." *Experimental Economics*, 4, 107-124.

ECKEL, C. and WILSON, R. (2004) "Is trust a risky decision?" *Journal of Economic Behavior & Organization*, 55, 447-465.

FALK, A., FEHR, E. and FISCHBACHER, U. (2005) "Driving Forces Behind Informal Sanctions," *Econometrica*, 73(6), 2017-2030.

FALKINGER, J, FEHR, E., GÄCHTER, S. and WINTER-EBNER, R. (2000) "A Simple Mechanism for the Efficient Provision of Public Goods – Experimental Evidence," *American Economic Review*, 90, 247-264.

FEHR, E. and FISCHBACHER, U. (2004) "Third Party Punishment and Social Norms," *Evolution and Human Behavior*, 25, 63-87.

FEHR, E. and GÄCHTER, S. (2000) "Cooperation and Punishment in Public Goods Experiments," *The American Economic Review*, 90, 980-994.

FEHR, E. and GÄCHTER, S. (2002) "Altruistic Punishment in Humans," *Nature*, 415, 137-140.

FEHR, E. and SCHMIDT, K. (1999) "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics,* 114, 817-868.

GÄCHTER, S., RENNER, E. and SEFTON, M. (2008) "The Long-Run Benefits of Punishment," *Science 322*, 1510.

GANGADHARAN, L. and NIKIFORAKIS, N. (2009) "Does the Size of the Action Set Matter for Cooperation?" *Economics Letters*, 104, 115-117.

HOLT, C. A. and LAURY, S. K. (2002) "Risk Aversion and Incentive Effects," *The American Economic Review*, 92, 1644-1655.

ISAAC, R. M. and JAMES, D. (2000) "Just Who Are You Calling Risk Averse," *Journal of Risk and Uncertainty*, 20(2), 177-87.

NIKIFORAKIS, N. and NORMAN, H.-T. (2008) "A Comparative Statics Analysis of Punishment in Public-Good Experiments," *Experimental Economics*, 11, 358-369.

OSTROM, E., WALKER, J. and GARDNER, R. (1992) "Covenants with and without a Sword: Self-Governance is Possible," *American Political Science Review*, 86(2), 404-417.

ROTH, A. (1988) "Laboratory Experimentation in Economics: A Methodological Overview," *The Economic Journal*, 98, 974-1031.

HAIGNER, S., KOCHER, M., and SUTTER, M. (forthcoming) "Choosing the stick or the carrot? Endogenous institutional choice in social dilemma situations," *Review of Economic Studies.*

WALKER, J. M. and HALLORAN, M. A. (2004) "Rewards and Sanctions and the Provision of Public Goods in One-Shot Settings," *Experimental Economics*, 7, 235–247

Appendix

Translation from German (column player version). The decision forms and the questionnaire were printed on separate sheets. The original instructions are available from the authors upon request.

1. GENERAL INSTRUCTIONS

INSTRUCTIONS

No talking: Now that the experiment has begun, we ask that you do not talk or communicate any longer with each other. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Monetary payments: The experiment will consist of two stages and will be followed by a separate decision problem for which you will get paid as well. The amount of money you make will depend on the choices made (as described below). Each participant will receive a lump sum payment of **5 Euro.** [Punishment: This one-off payment can be used to pay for eventual losses.] Your earnings will be paid to you in cash individually and privately at the end of the session.

Matching: During the session you will be matched with another person. However, no participant will ever know the identity of the person he or she is matched with.

Roles: In the experiment a "row" player is always randomly matched with a "column" player. You are the **column player**.

DECISION [Punishment: INSTRUCTIONS – STAGE 1]

Your decision [Punishment: Your decision in Stage 1]: On the DECISION FORM [Punishment: DECISION FORM 1] you will see a payoff table. The row player decides between Top and Bottom rows, and the column player decides between the Left and Right columns. The intersection of the designated row and column determines which part of the payoff matrix is relevant (Top Left, Top Right, Bottom Left, Bottom Right) and thus determines the earnings for each person. In each cell, the row player's payoff is shown first and the column player's payoff is shown second. Your payoff is printed in bold.

After you have made the decision, we will collect the decision forms and inform you about the decision of your matched row player.

DECISION FORM [Punishment: DECISION FORM – STAGE 1]

	Left	**Right**
Top	$5, **$5**	$0, **$8**
Bottom	$8, **$0**	$2, **$2**

Please **circle** either the **Left** or the **Right** column.

[Only in Punishment] INSTRUCTIONS – STAGE 2

Your Decision in Stage 2: After learning the other player's decision in Stage 1, you can **decrease** the other player's payoff by 2 Euro at the cost of 1 Euro to you.

If you decide to **decrease** the other player's payoff, you will **circle** the words "**I want to decrease the other player's payoff.**"

If you decide to **not change** the other player's payoff, you will **circle** the words "**I do not want to change the other player's payoff.**"

The other player can also decrease your payoff or leave it unchanged.

[Only in Punishment] DECISION FORM – STAGE 2

Do you want to decrease the other player's payoff by 2 Euro at the cost of 1 Euro to you? Please circle.

I want to decrease the other player's payoff.

OR

I do not want to change the other player's payoff.

2. INSTRUCTIONS RISK ATTITUDE ELICITATION

INSTRUCTIONS

The next page shows ten decision questions. Each decision is a paired choice between "Option A" and "Option B."

You will make ten choices and record these in the box to the left of the option. That is, if you prefer option A to option B, you will mark an X by option A. Only one of the ten decisions will be used in the end to determine your earnings for this part of the experiment.

A ten-sided die will be used to determine payoffs; the faces are numbered from 1 to 10 (the "0" face of the die will serve as 10.) After you have made all of your choices, we will throw this die twice, once to select one of the ten decisions to be used, and a second time to determine what your payoff is for the option you chose, A or B, for the particular decision selected. Even though you will make ten decisions, only one of these will end up affecting your earnings, but you will not know in advance which decision will be used. Obviously, each decision has an equal chance of being used in the end.

Now, please look at Decision 1 at the top. Option A pays 2.00€ if the throw of the ten sided die is 1, and it pays 1.60€ if the throw is 2-10. Option B yields 3.85€ if the throw of the die is 1, and it pays 0.10€ if the throw is 2-10. The other decisions are similar, except that as you move down the table, the chances of the higher payoff for each option increase. In fact, for Decision 10 in the bottom row, the die will not be needed since each option pays the highest payoff for sure, so your choice here is between 2.00€ or 3.85€.

To summarize, you will make ten choices: for each decision row you will have to choose between Option A and Option B. You may choose A for some decision rows and B for other rows, and you may change your decisions and make them in any order.

To determine the payoffs from this task we will throw the ten-sided die to select which of the ten Decisions will be used. Then we will throw the die again to determine the money earnings for the Option you chose for that Decision.

So now please look at the empty boxes on the record sheet. You will have to mark an X in one and only one of the boxes in each row, depending whether you prefer option A or option B. Then the die throw will determine which of the ten decisions is going to count. We will look at the decision that you made for the one that counts, and circle it, before throwing the die again to determine your earnings for this part.

DECISION FORM

Option A		Option B
☐ 2.00€ with probability of 1/10, 1.60€ with probability of 9/10	*OR*	☐ 3.85€ with probability of 1/10, 0.10€ with probability of 9/10
☐ 2.00€with probability of 2/10, 1.60€ with probability of 8/10	*OR*	☐ 3.85€ with probability of 2/10, 0.10€ with probability of 8/10
☐ 2.00€ with probability of 3/10, 1.60€ with probability of 7/10	*OR*	☐ 3.85€ with probability of 3/10, 0.10€ with probability of 7/10
☐ 2.00€ with probability of 4/10, 1.60€ with probability of 6/10	*OR*	☐ 3.85€ with probability of 4/10, 0.10€ with probability of 6/10
☐ 2.00€ with probability of 5/10, 1.60€ with probability of 5/10	*OR*	☐ 3.85€ with probability of 5/10, 0.10€ with probability of 5/10
☐ 2.00€ with probability of 6/10, 1.60€ with probability of 4/10	*OR*	☐ 3.85€ with probability of 6/10, 0.10€ with probability of 4/10
☐ 2.00€ with probability of 7/10, 1.60€ with probability of 3/10	*OR*	☐ 3.85€ with probability of 7/10, 0.10€ with probability of 3/10
☐ 2.00€ with probability of 8/10, 1.60€ with probability of 2/10	*OR*	☐ 3.85€ with probability of 8/10, 0.10€ with probability of 2/10
☐ 2.00€ with probability of 9/10, 1.60€ with probability of 1/10	*OR*	☐ 3.85€ with probability of 9/10, 0.10€ with probability of 1/10
☐ 2.00€ with probability of 10/10, 1.60€ with probability of 0/10	*OR*	☐ 3.85€ with probability of 10/10, 0.10€ with probability of 0/10

QUESTIONNAIRE

Thank you for participating in the experiment. Finally, please answer the following questions. Your answers will have no impact on your final payoff.

1. If you could decide how much to destroy of the other player's payoff how much would you like to destroy?
2. How much of your own payoff would you be willing to pay for it?
3. How old are you?
4. What is your gender?
5. What is your major?
6. In which country were you born/raised?