University of Heidelberg

Department of Economics

Discussion Paper Series | No. 501

Taking punishment into your own hands: An experiment on the motivation underlying punishment

Peter Duersch and Julia Müller

May 2010

Taking punishment into your own hands: An experiment on the motivation underlying punishment*

Peter Duersch† Julia Müller‡

May 25, 2010

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.

Keywords: personal punishment, real effort task, experiment, auction, desire to win.

JEL-Classifications: C92, D03.

*We thank for valuable comments: Maroŝ Servátka, Jörg Oechssler, Dirk Engelmann and Wolfgang Hoechtl. Financial support by the DFG through SFB 504 is gratefully acknowledged.

†Department of Economics, University of Heidelberg, email: peter.duersch@awi.uni-heidelberg.de

‡Department of Economics, University of Heidelberg, email: julia.mueller@awi.uni-heidelberg.de

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 com-

[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".

mon 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 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.

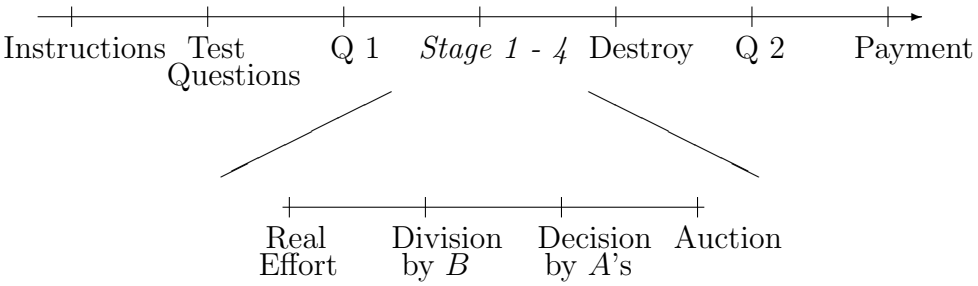

Figure 1: Timing

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

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

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.

The auction winner was informed about the second highest bid he had to

[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.

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.

[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.

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 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.

[7] In the experiment all subjects A completed the real effort task (compare section 5).
[8] See table III in Fehr and Schmidt (1999).

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.*

(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

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

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 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.

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. Howewer, the difference is not significant according to a MW test.[10]

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

[10] We also ran probit regressions for the decision to punish and bidding a positive amount,

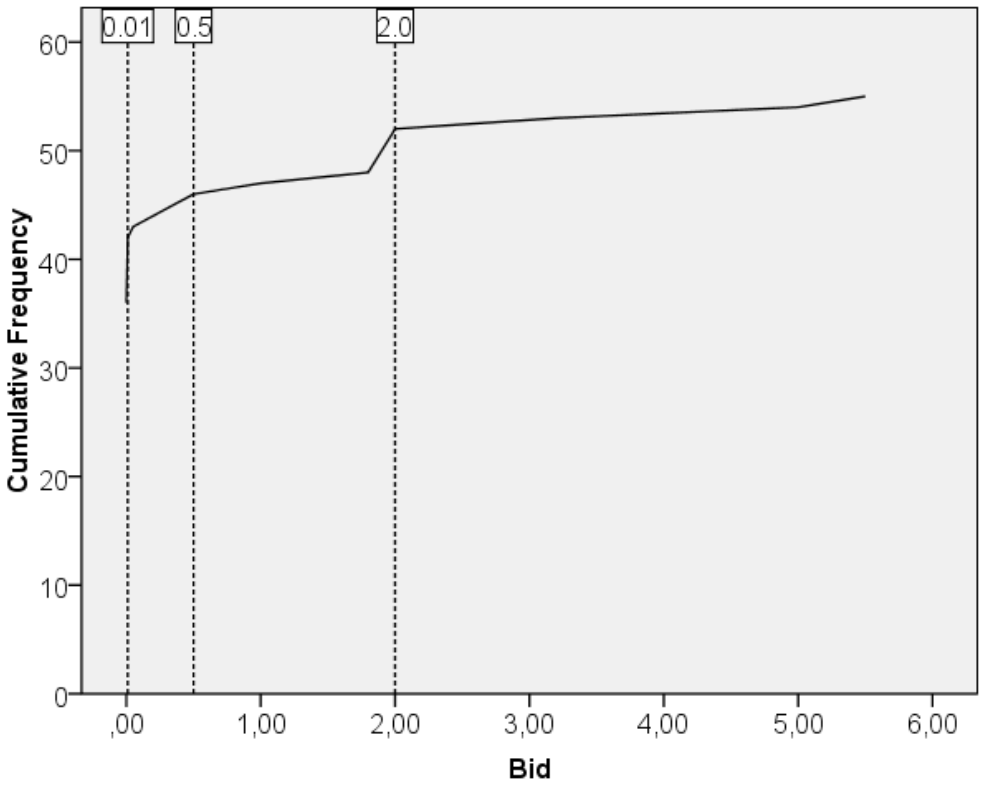

Figure 2: Cumulative distribution of bids

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

but we find our demographics and questionnaire data mostly unsignificant.

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

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.

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]

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

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 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.

[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

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).

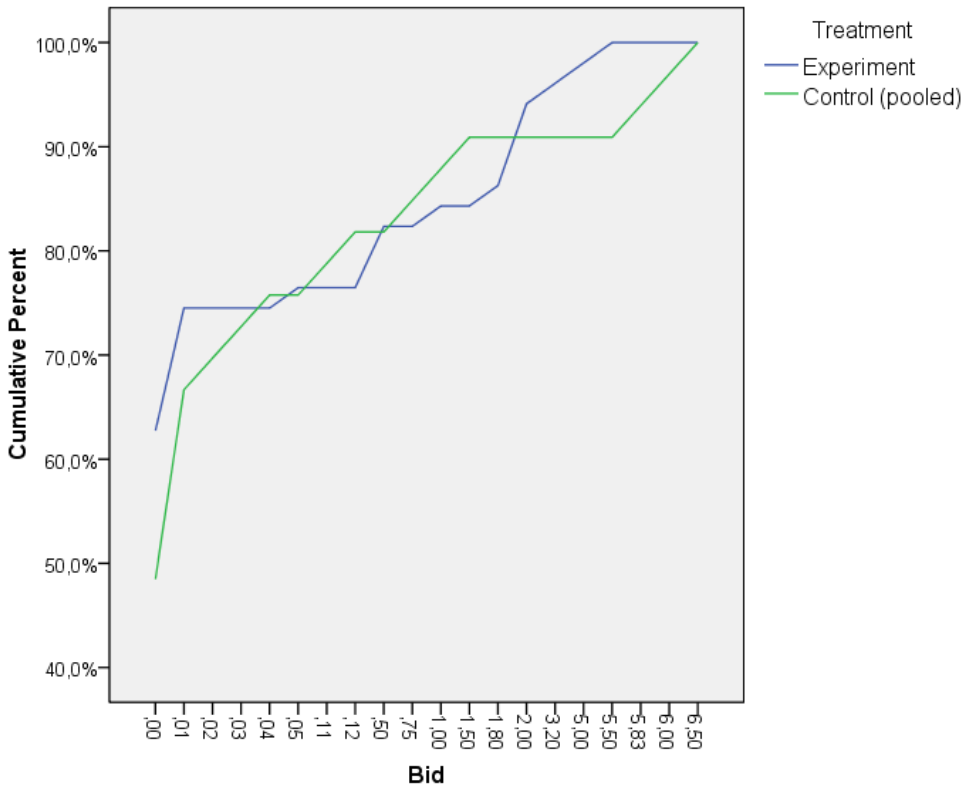

Figure 3: Cumulative distribution of bids by treatment

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 nonmonetary 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

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

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

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?