



TI 2006-104/1

Tinbergen Institute Discussion Paper

# Does Stake Size matter for Cooperation and Punishment?

*Martin G. Kocher<sup>1</sup>*

*Peter Martinsson<sup>2</sup>*

*Martine Visser<sup>2</sup>*

<sup>1</sup> CREED, Universiteit van Amsterdam, University of Innsbruck, and Tinbergen Institute;

<sup>2</sup> Göteborg University, Sweden.

**Tinbergen Institute**

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

**Tinbergen Institute Amsterdam**

Roetersstraat 31

1018 WB Amsterdam

The Netherlands

Tel.: +31(0)20 551 3500

Fax: +31(0)20 551 3555

**Tinbergen Institute Rotterdam**

Burg. Oudlaan 50

3062 PA Rotterdam

The Netherlands

Tel.: +31(0)10 408 8900

Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at  
<http://www.tinbergen.nl>.

# Does stake size matter for cooperation and punishment?\*

*Martin Kocher*<sup>#</sup>

CREED, University of Amsterdam, Netherlands  
Department of Economics, University of Innsbruck, Austria

*Peter Martinsson*

Department of Economics, Göteborg University, Sweden

*Martine Visser*

Department of Economics, Göteborg University, Sweden

This version: 2006-11-20

## **Abstract**

The effects of stake size on cooperation and punishment are investigated using a public goods experiment. We find that an increase in stake size does neither significantly affect cooperation nor, interestingly, the level of punishment

**Key words:** experiment, public goods, punishment, stake size

**JEL classification:** C72, C91, H41

\* We would like to thank Martin Dufwenberg, Olof Johansson-Stenman and participants of the ESA World Meeting 2006 in Atlanta for helpful comments as well as Benedikt Herrmann for suggestions regarding the research design. We are also indebted to the Western Cape Department of Education in South Africa for facilitating this research. Financial assistance from the Swedish International Development Cooperation Agency (SIDA) and from Jubileumsfonden at Göteborg University is gratefully acknowledged. Kocher also acknowledges financial support from the ENABLE Project under the European Union 6th Framework Program. The paper reflects the views of the authors, and the European Community is not liable for any use that may be made of the information contained herein.

<sup>#</sup> Corresponding author. Martin Kocher, CREED, University of Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, The Netherlands. E-mail: martin.kocher@uibk.ac.at. Ph: +31 (0)20 525 4250. Fax: +31 (0)20 525 5283.

# 1. Introduction

While some experimental economists argue that one of the concerns with data from the laboratory – the effect of stake size – has largely been put to rest, others provide evidence that this issue is still largely in dispute. Several studies that have explicitly tested for stake effects have found that an increase in monetary stakes does not significantly affect the average behavior of decision-makers. It could, however, reduce the variance of subjects' behavior in certain environments (e.g., Camerer, 2003; Camerer and Hogarth, 1999; Smith and Walker, 1993). Further evidence in line with this general conclusion has been provided, for instance, by Cameron (1999), Carpenter et al. (2005) as well as Slonim and Roth (1998) for dictator and ultimatum games, and Johansson-Stenman et al. (2005) for the trust game.<sup>1</sup> In contrast, critical remarks and counter-evidence can be found in Parco et al. (2002), Slonim and Roth (1998) as well as List and Levitt (2006). Thus, it would be premature to conclude that the effects of stake size should be neglected in experimental economics in general

The objective of this paper is to test whether stake size has an impact on cooperation and sanctioning behavior. We study both a standard linear public goods game and a public goods game that is augmented by a punishment stage after the contribution stage (Fehr and Gächter, 2000). Punishment takes the form of informal individual sanctions and is costly both to the punisher and the punished.

To the best of our knowledge, we are the first to analyze the effects of stake size on both voluntary contributions to a public good and on punishment. Most existing experimental evidence on stake size refers to bargaining games. Since social dilemmas are almost ubiquitous in everyday life, and since they often involve an option to sanction other decision-makers, we think that this is a worthwhile endeavor.

We are aware of only one other paper that explicitly deals with a test of stakes on the private provision of a public good. Marwell and Ames (1980) report that people invest less money in the public good when stakes are higher. Their finding, however, is confounded by an experimenter effect and they do not take punishment into account.

---

<sup>1</sup> In another context, decision-making under risk, Holt and Laury (2002) find no significant effect of stake size on risk attitudes.

## 2. The public good and our experimental design

Let  $I = \{1, 2, \dots, n\}$  denote a group of  $n$  subjects who interact only once in a one-shot simultaneous public goods game *without punishment*. Individual  $i \in I$  receives an endowment  $E$ , which can be allocated either to a private good or to a public good. The voluntary contribution of individual  $i$  to the public good,  $c_i$ , must satisfy  $0 \leq c_i \leq E$ . Let  $C$  denote the sum of all group members' contributions (i.e.  $C = \sum_{j=1}^n c_j$ ). Individual member  $i$ 's payoff from her contribution is given by

$$p_i = E - c_i + gC \quad (1)$$

The marginal per capita return (MPCR) from investing into this standard linear public good is denoted as  $g$ , which satisfies  $0 < g < 1 < ng$ , meaning that the self-interested choice and the social optimal one are in conflict.

In the public goods game *with punishment*, subjects can punish other group members individually, after they have received information on contribution levels of their group members. The game is now a two-stage game with a simultaneous contribution stage followed by a simultaneous punishment stage. Employing the punishment technology used by Gächter and Herrmann (2006), the unit cost per punishment point is 1, and one unit of punishment results in a deduction of 3 units in terms of payoff for the member who receives the punishment.<sup>2</sup> Each subject can assign a maximum of 10 punishment points to any other member in his or her group.

Taking into account the monetary consequences of the punishment stage yields the following payoff function for member  $i$ :

$$p_i = E - c_i + gC - 3 \sum_{k \neq i} p_{ik} - \sum_{h \neq i} p_{hi} \quad (2)$$

where  $p_{hi}$  is the cost of punishment by member  $h$  to member  $i$ , and  $p_{ik}$  is the deduction in terms of payoff as a consequence of punishment points from member  $i$  to member  $k$ .

Assuming rationality and selfishness, we should neither observe any voluntary contributions in both games nor any assignment of punishment points in the public goods game with punishment according to the subgame perfect equilibrium. Research shows, however, that subjects punish each other both in one-shot experiments as well as in multi-

---

<sup>2</sup> Previous experiments have shown that punishment behavior follows the law of demand, i.e. the quantity of punishment declines with a rising price of punishment (Anderson and Putterman, 2006; Carpenter, 2006).

period experiments with stranger matching (e.g., Fehr and Gächter, 2000; Gächter and Herrmann, 2006; Gächter et al., 2004).

The parameters in our experimental sessions were set up as follows: group size  $n = 4$ , endowment of  $E = 20$  Guilders (the experimental currency unit) and MPCR  $g = 0.5$ .

Each experimental session consisted of two parts. In *Part I*, each subject had to indicate his or her preferred contribution in the one-shot public goods game without punishment. After the decision and without any feedback on the results from this first part – to avoid order effects – subjects received experimental instructions<sup>3</sup> on *Part II*. Part II consisted of a public goods experiment that was augmented by a punishment stage after the contribution stage, according to the description above.<sup>4</sup> It was common knowledge that the group composition in this second part was different from the first part to rule out any reputation motives, and it was also announced that group members would remain anonymous.

Our experiment was conducted with 120 high school students (with an average age of 15.9 years) in Cape Town, South Africa, using paper and pen. In order to test for stake effects, the sample is divided into two groups: one with a low stake size treatment (*LOW*) and one with a high stake size treatment (*HIGH*). Most experiments identified in Camerer and Hogarth (1999) – but also more recently conducted stake experiments – use scale factors between 2 to 10 when testing for the effects of stake size.<sup>5</sup> Except for different conversion rates, sessions and instructions in *LOW* are identical to those in *HIGH*. In *LOW*, 1 Guilder is exchanged for 1.5 South African Rand (ZAR)<sup>6</sup>, whereas in *HIGH*, 1 Guilder is worth 7.5 ZAR.

In *LOW*, the average income per hour was intended to be a bit higher than an average hourly salary, while average income per hour in *HIGH* should correspond approximately to a daily salary. Note that hourly wages for casual and unschooled labor in South Africa ranges from 5 to 25 ZAR (Department of Labour, South Africa, 2006). Actually, on average, subjects earned 65.45 ZAR in *LOW* and 338.56 ZAR in *HIGH*, and sessions lasted slightly more than two hours. Thus, already the stakes in *LOW* were salient.

---

<sup>3</sup> The complete experimental instructions for both parts can be found at [URL will be provided for publication]. The instructions were phrased in neutral terms. Participants were not instructed to maximize their earnings and no references to any specific strategies were made.

<sup>4</sup> In Part II of the experiment, the impact of punishment was capped at the amount earned in the contribution stage. However, the punisher incurred the costs of punishing even if it resulted in a loss in Part II. This procedure was common knowledge among subjects, and there was actually no case of a loss in the experiment.

<sup>5</sup> There are, however, exceptions, for example a factor of 25 used in Slonim and Roth (1998), 20 in Johansson-Stenman et al. (2005) and up to 50 in Cameron (1999). These three experiments were conducted in relatively poor countries (Slovakia, Bangladesh and Indonesia, respectively).

<sup>6</sup> The exchange rate at the time of the experiment was 6.10 ZAR = 1 USD.

Both treatments were run at the same school, and the treatments were scheduled in over-lapping succession to avoid contagious effects by word-of-mouth communication. Participants were randomly allocated into the two treatments, decisions were taken anonymously, and communication among participants was prohibited. In addition to instructions detailing each step of the experiment, we used a number of quiz questions to ensure that everybody understood the task completely before participants made their choices. Final payment of experimental profits was private and in cash.

### 3. Results

In Table 1, we show mean contribution levels to the public good in Part I and Part II, separately for the *LOW* and the *HIGH* treatment. As can easily be seen in the table, there are only small differences between the two treatments. In Part I, the average contribution level was 34.4% in *LOW* and 32.9% in *HIGH*, while in Part II it was 41.2% and 40.9% for *LOW* and *HIGH*, respectively. Punishment in Part II was used by 25% of participants in *LOW* and by 17% of participants in *HIGH*. The average amount of punishment points awarded to another group member was rather low, however, namely 0.49 in *LOW* compared to 0.31 in *HIGH*. For positive levels of punishment, i.e.  $p_{hi} > 0$ , average punishment was 1.96 in *LOW* and 1.85 in *HIGH*, respectively. We apply Kolmogorov-Smirnov two-sample tests and Mann-Whitney-U tests to test the null hypotheses that (i) the contributions to the public good, and (ii) punishment points awarded in the two treatments come from populations with the same distribution and have equal means. We cannot reject these null hypotheses at any conventional levels.<sup>7</sup>

>>> TABLE 1

Figure 1 displays the direction of punishment dependent on the difference between the contribution of the punished player and one's own contribution. It shows that punishment in both treatments is predominately directed towards free-riders by high contributors (on the left-hand side of the graph); however, there are also a small number of free-riders that punish contributors. This is in line with the usual pattern observed in other studies on punishment. The only remarkable difference between our two treatments is a stronger

---

<sup>7</sup> Standard deviations are a little bit smaller in *HIGH* than in *LOW* but the difference is also far from being significant.

punishment reaction for large negative deviations in LOW than in HIGH, although the overall absolute level of punishment is still relatively small. The spike in the HIGH treatment at the interval (8,14] is due to only one subject, who probably made a mistake, choosing the maximum of ten punishment points.<sup>8</sup>

We do not observe any significant effects of important socio-economic variables like gender or age on contributions and punishment or any interaction effect of these variables with stake size.<sup>9</sup>

>>> FIGURE 1

## 4. Conclusion

This paper analyzes the effects of a stake size variation in a one-shot public goods experiment with and without punishment. The substantial increase in stakes does neither have a significant effect on the mean nor on the variance of the level of contributions in our experiment. It also has no significant effect on punishment. Our results suggest that findings of public goods experiments with standard laboratory stakes can be extrapolated to situations with considerable stakes. Evidence from a game show whose setup is related to the public goods game to a certain extent (Oberholzer-Gee and Waldfogel, 2003) shows that our conclusion is also likely to extend to exceptionally high stakes.

---

<sup>8</sup> Excluding this subject would not result in an overall significant difference in punishment behavior across the two treatments.

<sup>9</sup> Regressions are available from the authors upon request.



## References

- Anderson, C.M. and L. Putterman, 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.
- Camerer, C.F. , 2003, *Behavioral Game Theory*, Princeton University Press, New York.
- Camerer, C.F. and R.M. Hogarth, 1999, The effects of financial incentives in experiments: A review and capital-labor-production framework, *Journal of Risk and Uncertainty* 19, 7-42.
- Cameron, L.A., 1999, Raising the stakes in the ultimatum game: Experimental evidence from Indonesia, *Economic Inquiry* 37, 47-59.
- Carpenter, J., E. Verhoogen and S. Burks, 2005, The effect of stakes in distribution experiments, *Economics Letters* 86, 393-398.
- Carpenter, J., 2006, The demand for punishment, *Journal of Economic Behavior and Organization*, forthcoming.
- Department of Labour, South Africa, 2006, Basic Conditions of Employment Amendment Act, 2002, <http://www.labour.gov.za/>.
- Fehr, E. and S. Gächter, 2000, Cooperation and punishment in public goods experiments, *American Economic Review* 90, 980-994.
- Gächter, S., B. Herrmann and C. Thöni, 2004, Trust, voluntary cooperation, and socio-economic background: Survey and experimental evidence, *Journal of Economic Behavior and Organization* 55, 505-531.
- Gächter, S. and B. Herrmann, 2006, The limits of self-governance in the presence of spite: Experimental evidence from urban and rural Russia, Working Paper 2006-13, University of Nottingham.
- Holt, C.A. and S.K. Laury, 2002, Risk aversion and incentive effects, *American Economic Review* 92, 1644-1655.
- Johansson-Stenman, M. Mahmud and P. Martinsson, 2005, Does stake size matter in trust games?, *Economics Letters* 88, 365-369.
- List, J.A. and S.D. Levitt, 2006, What do laboratory experiments tell us about the real world? Discussion Paper, University of Chicago.
- Marwell, G. and R.E. Ames, 1980, Experiments on the provision of public goods. Provision points, stakes, experience, and the free-rider problem, *American Journal of Sociology* 85, 926-937.

- Oberholzer-Gee, F. and J. Waldfogel, 2003, Social learning and coordination in high-stakes games: Evidence from friend or foe, Working Paper No. 2003-01, CREMA.
- Parco, J.E., A. Rapoport and W.E. Stein, 2002, Effects of financial incentives on the breakdown of mutual trust, *Psychological Science* 13, 292-297.
- Slonim, R. and A.E. Roth, 1998, Learning in high stakes ultimatum games: An experiment in the Slovak Republic, *Econometrica* 66, 569-596.
- Smith, V. and J. Walker, 1993, Rewards, experiences and decision costs in first price auctions, *Economic Inquiry* 31, 237-244.

Table 1. Levels of contribution to the public good (within parentheses: proportion contributed)

Treatments	Part I (without punishment)	Part II (with punishment)	Av. punish- ment points awarded	Proportion of punishers ( $p_{hi} > 0$ )	Av. punish- ment points awarded if ( $p_{hi} > 0$ )
LOW	6.88 (34.4%)	8.25 (41.2%)	0.49	0.25	1.96
HIGH	6.58 (32.9%)	8.18 (40.9%)	0.31	0.17	1.85
LOW vs. HIGH: p-values*	0.99/0.94	0.81/0.88	0.80/0.15		0.81/0.64
LOW vs. HIGH: p-values**				0.26	

Note. \* Based on Kolmogorov-Smirnov two-sample tests before slash and Mann-Whitney-U tests after slash (two-sided). \*\* Based on two-sample test of proportions.

Figure 1. Direction of punishment

