

Punishment without crime: a tale of competition and cooperation in Public Goods Games

Alexis Belianin

ICEF and BELab, NRU HSE, and IMEMO RAS
<http://epee.hse.ru>
icf-research@hse.ru

September 29, 2016
CCCP conference
29/09/16, Moscow, Russia

Public goods game

Problem statement

Punishment motives

Experiment

Results

- General trends

- Hypotheses

- Survey

- Reduced-form model

- Behavioural model

Conclusions

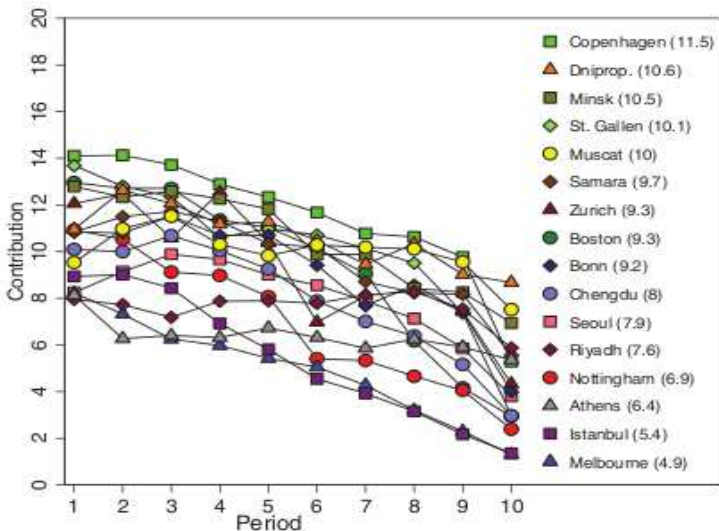
Public goods (PG) game with voluntary contribution mechanism (VCM)

- $n \geq 2$ players endowed with (normalized) 1 unit per period each.
- Each player i independently decides what fraction $c_i, 0 \leq c_i \leq 1$ she will contribute to the public good, retaining $1 - c_i$.
- Return from public good is $k \cdot \sum_i c_i = \alpha \bar{c}$, where $\bar{c} = \frac{\sum_i c_i}{n}$ and $\alpha = kn, k < 1 < kn$ is efficiency factor.

$$v_i = 1 - c_i + \alpha \bar{c} = 1 - c_i + k \cdot \sum_i c_i \quad (1)$$

The only Nash equilibrium is zero contribution, while social optimum is 100% contribution

PG with VCM: typical results (Herrmann, Gächter, Thöni, 2009)



Public goods game with VCM and punishment

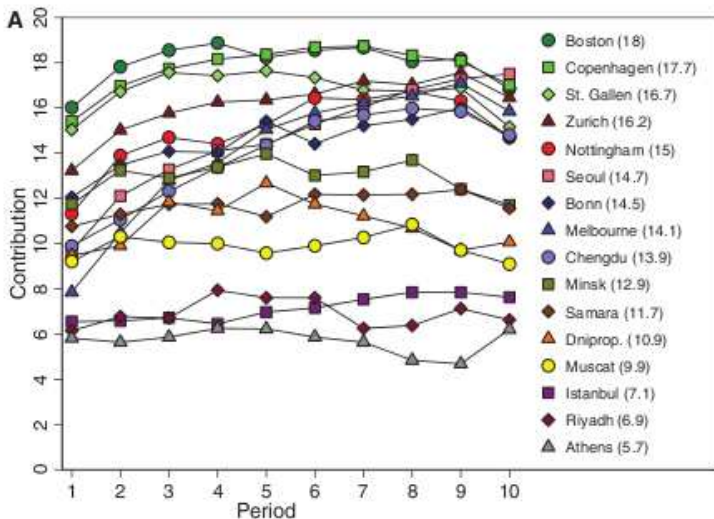
After the contribution stage, all players are informed about individual contributions, and can punish each other player j (not herself!) by p_{ij} units at a cost sp_{ij} units to themselves, where $s < 1$. Total payoff to player i is then

$$V_i(\mathbf{c}, \mathbf{P}) = v_i - s \sum_{j \neq i} p_{ij} - \sum_{j \neq i} p_{ji} \quad (2)$$

Punishments are known to increase the degree of cooperativeness, especially in with time and in partner treatments.

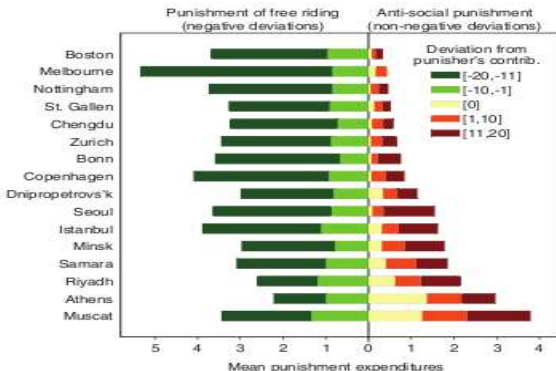
Mechanism: punishment (threaten, expression of disapproval) of those who free-ride boosts up cooperativeness.

PG with VCM: typical results (Herrmann, Gächter, Thöni, 2008)



Antisocial punishment (Herrmann, Gächter, Thöni, 2008)

Sometimes players punish not only those who contributed less, (free-riders — *prosocial* punishment), but also those who contributed more than they did (*spiteful*, or antisocial punishment)



Middle East and Eastern Europe are world leaders in antisocial attitudes...

...or are they?

- What are the reasons why spiteful punishment is more widespread in some parts of the world rather than in others?

...or are they?

- What are the reasons why spiteful punishment is more widespread in some parts of the world rather than in others?
- More generally: what are the motives for antisocial and prosocial punishment behaviour?

...or are they?

- What are the reasons why spiteful punishment is more widespread in some parts of the world rather than in others?
- More generally: what are the motives for antisocial and prosocial punishment behaviour?
- We address these questions by triple means:
 1. Experimental design that reveals subjects' preferences to punish each other for particular reasons;
 2. Survey questionnaire soliciting verbal intentions of the subjects;
 3. Structural experimental model predicting types of punishers based on experimental data.

Summary of punishment motives

Availability — presense of punishment option implicates punishment

Summary of punishment motives

Availability — presence of punishment option implicates punishment

Tolerance — culturally-conditioned punishment as something customary and normal

Summary of punishment motives

Availability — presence of punishment option implicates punishment

Tolerance — culturally-conditioned punishment as something customary and normal

Preemption — punishing because one expect punishment from others

Summary of punishment motives

Availability — presence of punishment option implicates punishment

Tolerance — culturally-conditioned punishment as something customary and normal

Preemption — punishing because one expect punishment from others

Jealousy — culturally-conditioned punishment of those who 'show up' themselves

Summary of punishment motives

Availability — presence of punishment option implicates punishment

Tolerance — culturally-conditioned punishment as something customary and normal

Preemption — punishing because one expect punishment from others

Jealousy — culturally-conditioned punishment of those who 'show up' themselves

Competitiveness — punishment as an efficient way to improve own relative standing in the group

Summary of punishment motives

Availability — presense of punishment option implicates punishment

Tolerance — culturally-conditioned punishment as something customary and normal

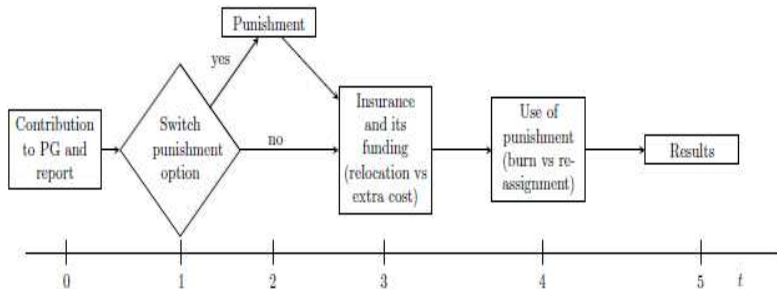
Preemption — punishing because one expect punishment from others

Jealousy — culturally-conditioned punishment of those who 'show up' themselves

Competitiveness — punishment as an efficient way to improve own relative standing in the group

Retaliation — punishment caused by negative feeling at what the others have contributed, be these generous or greedy

Flow of the experiment



Testing for the motives

Availability	Stage to switch on the punishment option
Tolerance	Unwillingness to insure at all
Preemption	Insurance by money relocation from punishment
Jealousy	Punishment alongside with insurance
Competitiveness	Burning money
Retaliaiton	Passing money to generous/greedy players

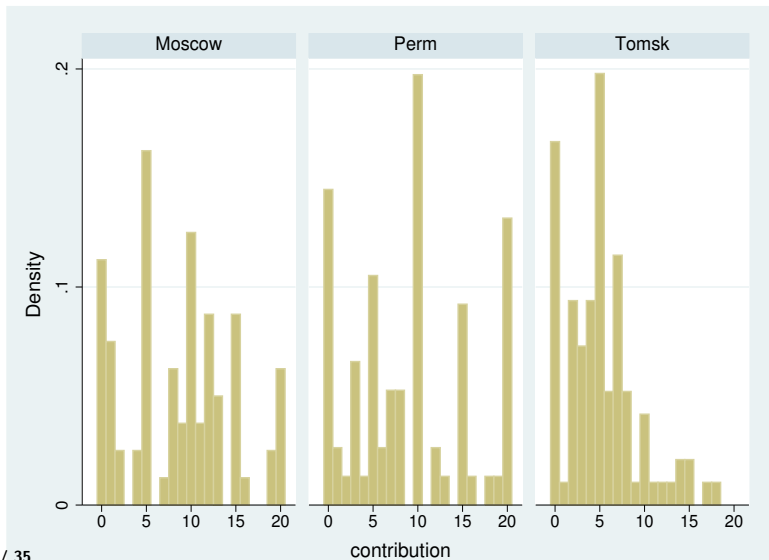
Design (after Gächter and Herrmann (2008))

- 2 single-shot games: VCM without punishment, followed by VCM with punishment (2 games altogether).
- Groups of $n = 4$ players, endowment 20, efficiency factor $k = 1.6$ ($\alpha = 0.4$) for all subjects.
- After each contributions stage, participants observe contributions and payoffs of all groupmates.
- Cost of punishment from 0 to 10 either low (0.1) or high (0.5).
- Preceding instructions with worked examples and exercises to check understanding.
- Ex ante intentions questionnaire other than oneself and the punished one, in proportion to their contributions.
- Post-punishment treatments introduced consecutively, so that subjects are not aware of the contents of the next one before completing the previous one.

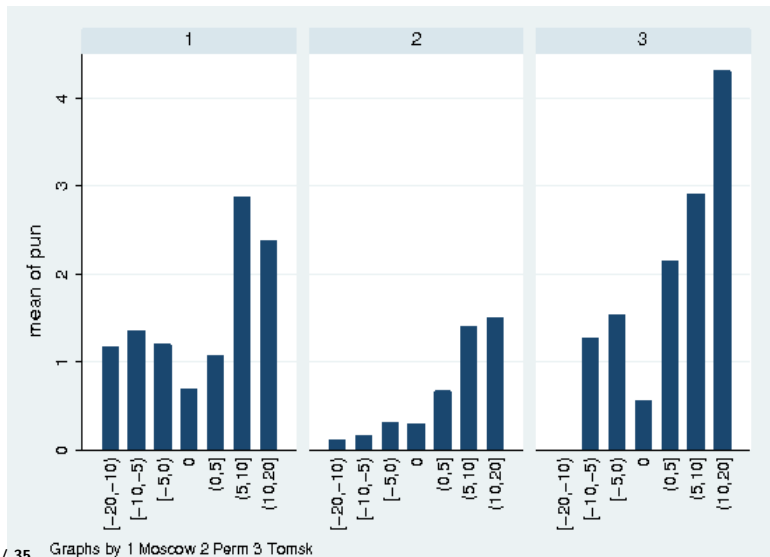
Design: details

- Intentions questionnaire asks for *planned* own contributions, the *due* average and *expected* average contributions in their group, and desired contribution level if the group average turns out to take discrete values of 0, 3, 6, 10, 14 and 17 units, evaluated by strategy method.
- One-shot game: public good without punishment followed by public good with punishment
- Participants: 364 full-time and part-time students from Moscow (193), Perm (76) and Tomsk (96). Average payoff — 320 RuR. (around 8 euro at the time of experiment).

Contributions: distribution



Punishments by differences in contributions

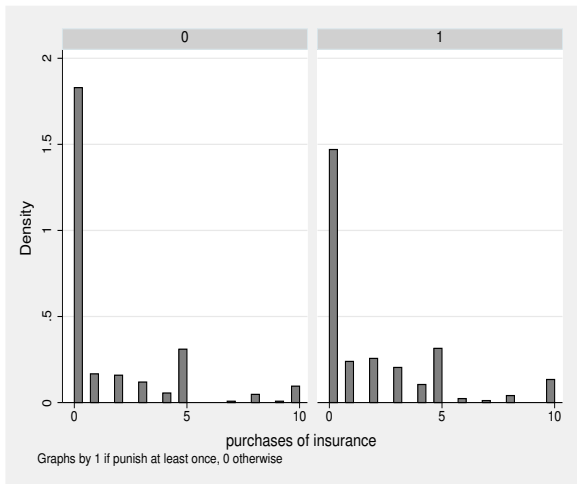


Availability

city	instances	wish share	punishment share
Moscow	768	.56	.27
Perm	304	.44	.17
Tomsk	384	.61	.28
Overall	1456	.55	.25

- Overall, 163 players did not wish to punish, and 201 did. However, 18 out of 201 players who intended to punish did not eventually punish anyone (!).
- 25% of all possible punishment instances took place, with 183 players punishing above zero at least once. This is in line with the previous experiments, so availability does not seem to be crucial.

Tolerance

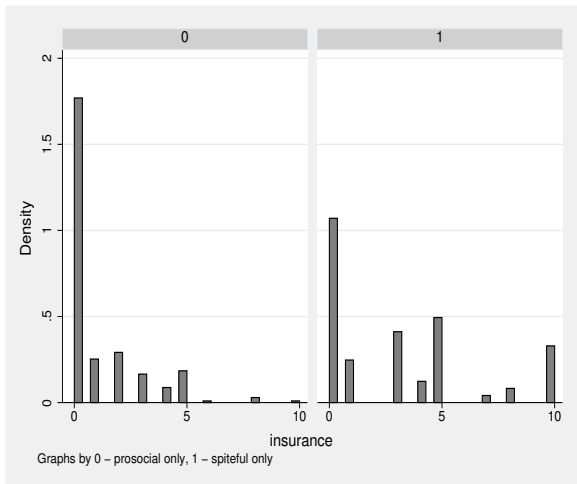


Tolerance: breakdown

	no punishment	punishment > 0	Total	prosocial	spiteful
no insurance	117	90	207	127	44
insurance > 0	61	96	157	111	69
Overall	178	186	364	238	113

- Instances of insurance are positive for both punishers and non-punishers, ruling out tolerance.
- Punishers insure more frequently than non-punishers, esp. when punishment is spiteful.

Preemption & Jealousy



Preemption & Jealousy: breakdown

Instances based on only prosocial/spiteful players

	redistribute (insure only)	extra money (insure and punish)	Total
prosocial	141	147	288
spite	50	18	68
Overall	191	165	356

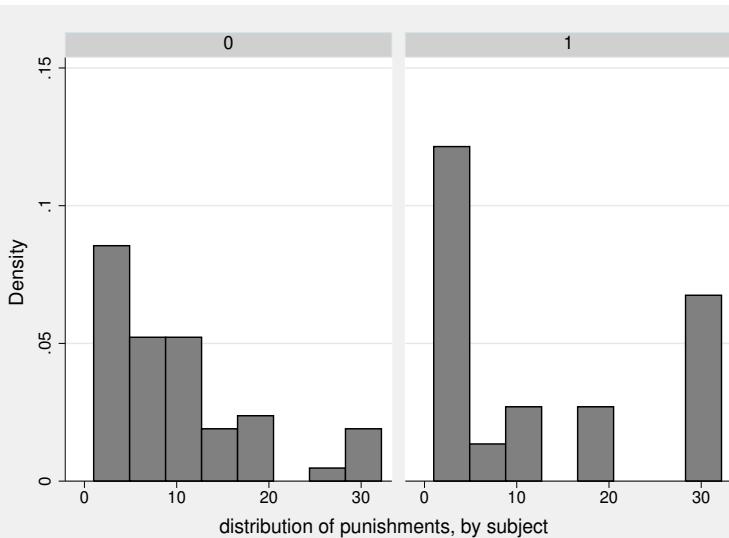
- Overall, 113 players punished only prosocially, and 27 (about 20%) only spitefully.
- Most spiteful punishers are preemptive (relocate money to insurance if they can), while prosocial punishers are both jealous and preemptive.

Competitiveness & retaliation

		burn	redistribute	Total
Breakdown table by subjects	prosocial	48	65	113
	spite	16	11	27
	Overall	64	76	140

- Overall, 140 players had to reassign something, of which about half decided to burn and to redistribute money.
- Spiteful punishers significantly more competitive (burn money), while prosocial punishers redistribute more often, esp. if partners were generous.

Distribution of punishment expenditures



Punishments: findings

confirmed: Mean frequency and size of spiteful punishments are compatible with those of the previous experiments.

Punishments: findings

- confirmed:** Mean frequency and size of spiteful punishments are compatible with those of the previous experiments.
- confirmed:** Mean punishment size and cost per punishment is the same for prosocial and spiteful punishments (similar rationality)

Punishments: findings

- confirmed:** Mean frequency and size of spiteful punishments are compatible with those of the previous experiments.
- confirmed:** Mean punishment size and cost per punishment is the same for prosocial and spiteful punishments (similar rationality)
- new!** Spiteful punishments are typically more serial than prosocial, and also punish more (on average, by 13, vs. 8.7 for prosocial ones)

Punishments: findings

- confirmed:** Mean frequency and size of spiteful punishments are compatible with those of the previous experiments.
- confirmed:** Mean punishment size and cost per punishment is the same for prosocial and spiteful punishments (similar rationality)
- new!** Spiteful punishments are typically more serial than prosocial, and also punish more (on average, by 13, vs. 8.7 for prosocial ones)
- new!** Spiteful punishers tend to be driven more by preemption and competition, while prosocial ones — by jealousy and retaliation.

Punishments: findings

- confirmed:** Mean frequency and size of spiteful punishments are compatible with those of the previous experiments.
- confirmed:** Mean punishment size and cost per punishment is the same for prosocial and spiteful punishments (similar rationality)
- new!** Spiteful punishments are typically more serial than prosocial, and also punish more (on average, by 13, vs. 8.7 for prosocial ones)
- new!** Spiteful punishers tend to be driven more by preemption and competition, while prosocial ones — by jealousy and retaliation.
- new!** Availability and tolerance do not seem to be crucial.

Survey questionnaire

Reasons for punishments		
Variable	Prosocial (N=121)	Spiteful (N=53)
Lower (than average) contribution	47.1	20.8
To stop them lowering our revenues	13.2	7.5
To gain more than they will	12.4	43.4
Afraid of them reducing my revenue	11.8	9.4
To equalize revenue within group	9.1	15.1
Intuitively/to experiment	7.5	1.9

Size determinants		
Variable	Prosocial (N=121)	Spiteful (N=50)
Inverse to their contribution	29.0	6.0
Maximal to the smallest contributor	18.5	8.0
To average out revenue	15.5	16.0
To put all revenues down to mine	11.5	—
Intuitively	8.7	14.0
Depending on my costs	6.8	—
Maximal to all	2.9	38.0
Minimal to all	1.9	8.0

Punishments factors: Tobit model estimates

Variable	Spiteful		Prosocial		Overall	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
<i>contr</i>	-0.35***	(0.12)			-0.17	(0.12)
<i>dcontr</i>			0.26***	(0.06)	0.51***	(0.15)
<i>rcontr</i>	-0.88***	(0.20)			-0.57***	(0.19)
<i>econtrx</i>	0.20*	(0.12)			-0.02	(0.10)
<i>econtra</i>			-0.11**	(0.05)	0.11	(0.07)
<i>condev</i>	0.28**	(0.11)			0.06	(0.09)
<i>tomsk</i>	3.25***	(1.21)			2.97***	(0.81)
<i>const</i>	4.34**	(1.65)	2.95***	(0.47)	-4.02***	(1.26)

*** — significant at 1%, ** — significant at 5%, * — significant at 10%

contr — c_j of punished

rcontr — $\Delta(c_i - \bar{c})$

econtrx — $\Delta(c_i - E c_i)$

condev — $\Delta(c_i - E c_i)$ at group mean

tomsk — dummy for Tomsk, cost 0.1

dcontr — $\Delta(c_i - c_j)$

econtra — $\Delta(c_i - E \bar{c})$

Punishment factors: interpretations

1. Many of the prosocial punishments are caused by **jealousy**: 1) differences in contributions and 2) over-contribution of the punisher relatively to her normative group standard. Secondary factor seems to be *retaliation*.
2. Most spiteful punishments are serial and **preemptive**, but some are 2) indiscriminate and largest for all partners, in line with **competitive** explanation.
3. None of the explanatory variables for one type of behaviour is significant as explanatory variable for the other

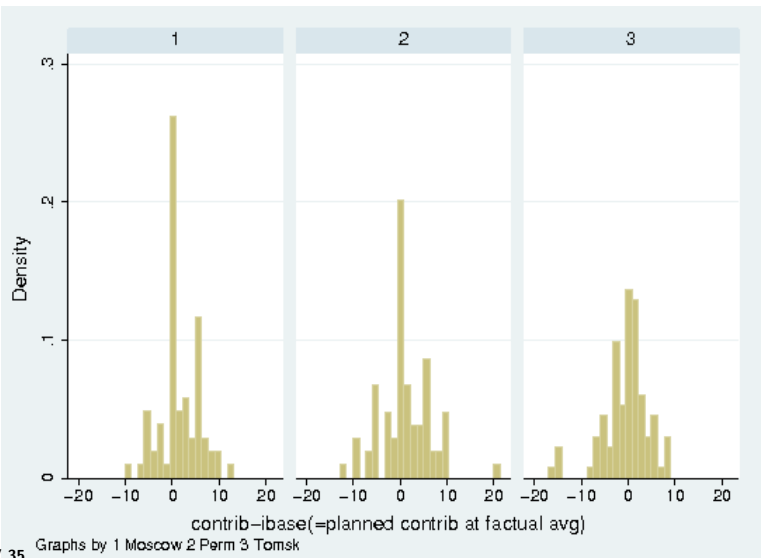
Behavioural model of punishment motives

$$u_i = V_i + \lambda_{1i} \frac{\sum_j \sum_k \gamma_k \varphi_{kij}}{p_{ij}} - \lambda_{2i} \sum_j \frac{E p_{ji}}{p_{ij}} - \pi \left[\lambda_{1i} \sum_j p_{ji} \left(\sum_k \gamma_k \varphi_{kij} \right) + \lambda_{2i} \sum_j E p_{ji} \right] \quad (3)$$

- V_i — material payoff,
- φ — retaliation (prosocial) / competition (spiteful) function of player i at player j ,
- $E p_{ji}$ — expectation of player i of punishment from player j , triggering jealousy (prosocial) and preemption (spiteful) motives.
- π — cost of punishment,
- λ_{1i} and λ_{2i} — individual-specific weights to retaliation/jealousy (prosocial) and competition/preemption for spiteful punishment motives

Maximizing (3) wrt punishment p_{ij} ,

Factual vs strategic form planned contributions



Model estimates

For prosocial punishment:

$$pun = \alpha + \lambda_1 \phi(prcontr + pcontr) + \lambda_2 \phi(pcons) + \varepsilon \quad (5)$$

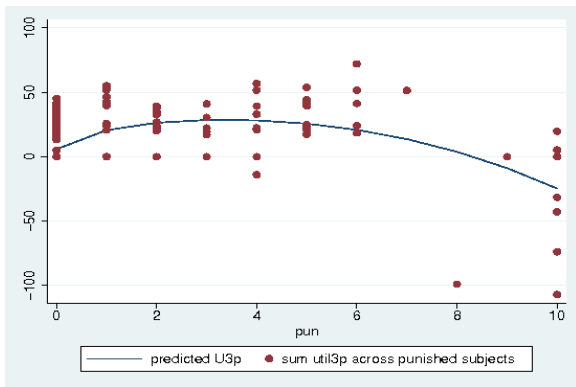
Weights are $\lambda_1^P = 0.207$, $\lambda_2^P = 0.793$, implying larger proportion of jealous players.

For spiteful punishment:

$$pun = \alpha + \lambda_1 \phi(pcondev) + \lambda_2 \phi(pcons) + \varepsilon \quad (6)$$

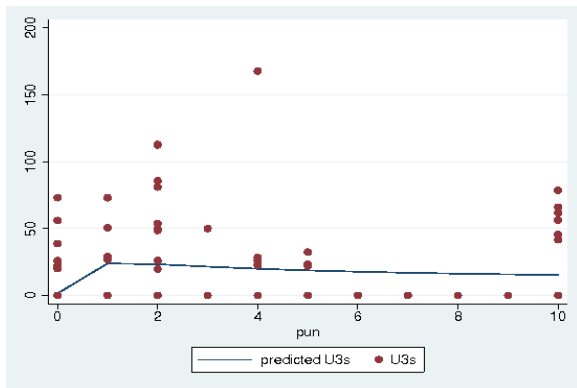
Weights $\lambda_1^P = 0.826$, $\lambda_2^P = 0.176$, imply larger proportion of retaliators.

Estimated utility for prosocial punishers



Inverse U-shape of utility vs. punishment size: at lower levels, larger punishments correspond to low utility of the punisher as they reflect their unhappiness with the social behaviour.

Estimated utility for spiteful punishers



U-shape graph with high dispersion at low punishment levels and large utility for those with extreme punishments.

Classification: the four punishment categories

- Active prosocial — retaliation (15%)** Punishments motivated by low contributions of the punished relative to the group standard. Believe they are on their right, punish by a lot (mean 9.78), and almost do not insure (mean 1.28).
- Passive prosocial — jealousy/retaliation (58%)** Appreciate fair behaviour, but unwilling/afraid of expression for the good cause, and/or cost concerned. Punishment is low (3.51), insurance yet lower (2.5).
- Passive spite — preemptive (17%)** Afraid of being exploited by the others. Both punishments (2.66) and insurance (2.5) are low.
- Active spite — competitive (12%)** Motivated by competitiveness, but also very afraid of preemption: use maximal punishments (10 in 100% cases) and insurance (7.38%).

Interpretations and extensions

- Punishment in PG context at least, should not always be interpreted as a revelation of dissatisfaction with contributions of the other players: there is a variety of competing explanations.
- These results suggest a multiplicity of principles on which 'punishment' behaviour may rest. In Russia, these were quite heterogeneous, while in Western Europe, for instance, 'spiteful' punishments are minor. Decomposition of punishment motives may be interesting and important for the diagnosis of the state of the respective societies.

Thank you!