Cooperation in social dilemmas

Martin Kocher
University of Munich,
University of Gothenburg, and CESifo Munich

March 2012
Queensland University of Technology

Motivation I

Radiohead, *In Rainbows*

Missing lighthouse?

Motivation II

- Cooperation is a ubiquitous phenomenon that has not only been studied in humans but also in many animals.
- It is, of course, especially interesting to study cooperation in contexts in which individual rationality (under selfishness assumptions) would dictate defection but collective rationality (efficiency) would require cooperation.
- Such situations have been coined social dilemmas (in sociology and psychology). In economics cooperation is usually analyzed in prisoner's dilemma games (binary) or in the context of the provision of public goods (public goods games).

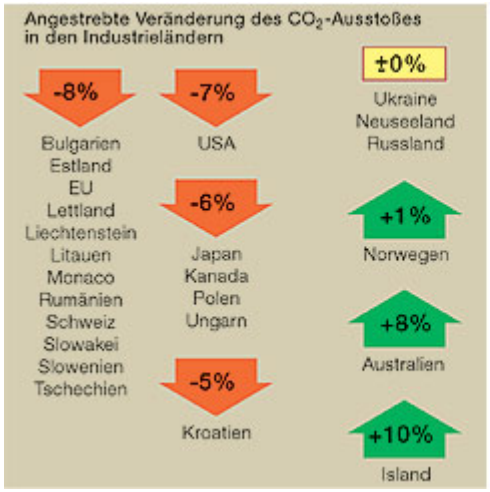
Motivation III

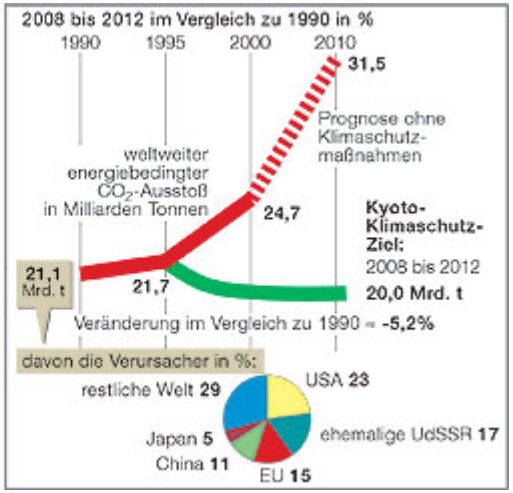
- Applications are so numerous that it is difficult not to forget important ones.
- Among the most important are: effort in work teams, provision of tangible or intangible public goods, the use of natural resources (the tragedy of the commons), contributions to charities, tax honesty, buyer-seller relationships, employer-employee relationships.
- Research on cooperation became clearly interdisciplinary lately, with many biologists, mathematicians, sociologists, psychologists and political scientists working (also jointly) to answer similar research questions.

Motivation IV

How to overcome social dilemma problems?

- Voluntary contribution
- Informal institutions
- Contracts (formal institutions)





The basic questions

- Why do people cooperate when it is individually optimal to defect?
- What explains the dynamics of cooperation?
- Which kind of institutions foster cooperation?
- How are they implemented?

Workhorses

- In the early days of experiments in sociology, psychology and economics the prisoner's dilemma game was used to study cooperation.
- Now, most studies use a simultaneous linear public goods game (the voluntary contribution mechanism = VCM).
- Bear in mind other mechanisms such as step-level public goods (provision point mechanisms), non-linear public goods, dynamic public goods, sequential contributions etc.

Summary public goods experiments I

Why do people contribute to public goods when it is individually rational not to?

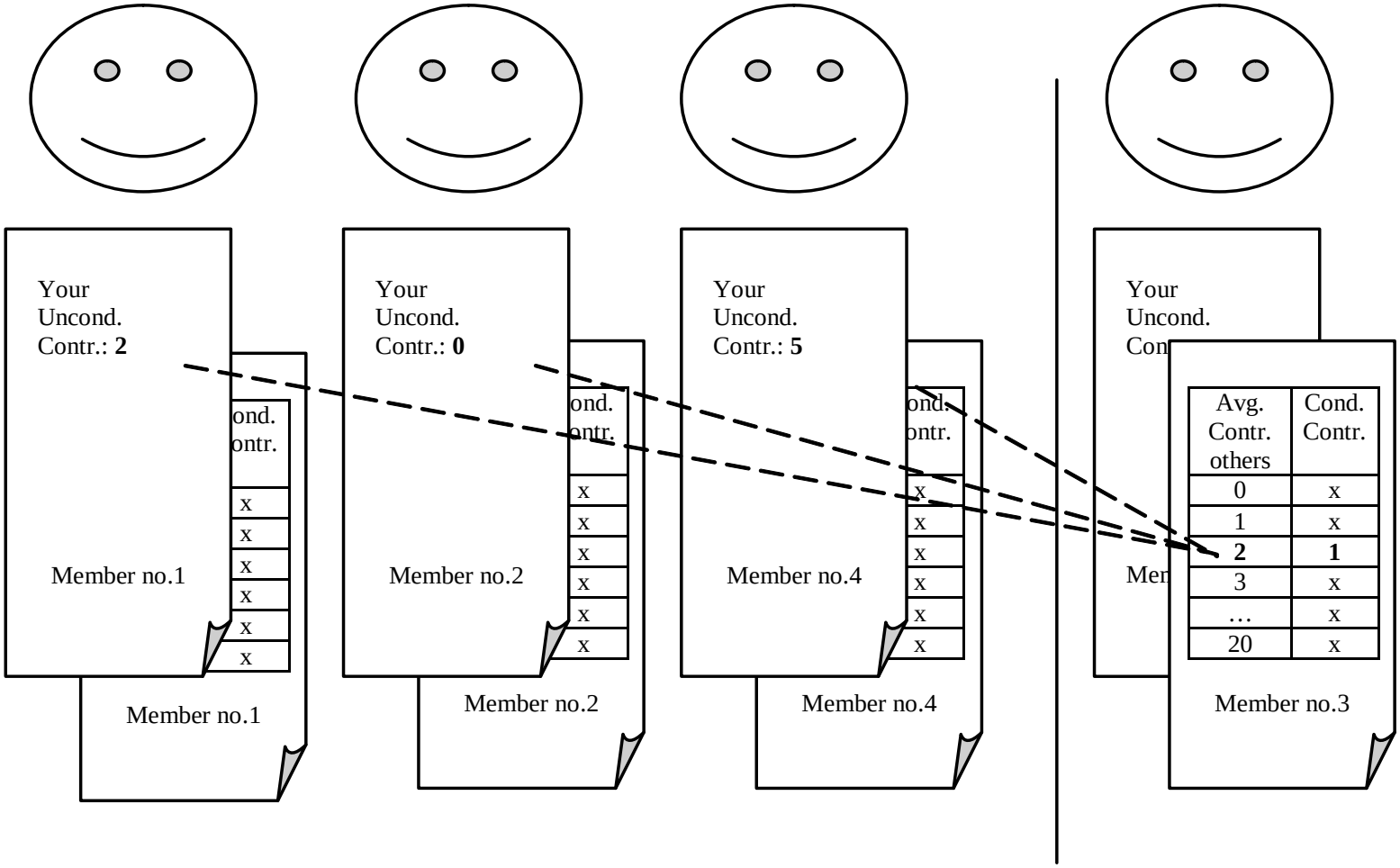
- Altruism, warm glow
- Fairness
- Conditional cooperation, reciprocity
- Confusion, mistakes

Palfrey and Prisbrey (1997, AER) and Kurzban and Houser (2002, AER) versus Fischbacher, Fehr and Gächter (2001, Econ Lett) as well as some discussion papers.

Fischbacher et al. (2001) - Idea

- The basic idea is to classify types of players.
- This is achieved by using a variant of the strategy vector method.
- One wants to know how much somebody is willing to contribute to the public good if she knows the contribution level of other players.
- Can also be done over time (Keser and van Winden, 2000) or by eliciting beliefs – but not as nicely.
- Base game: Standard linear public goods game.
- The resulting conditional contribution schedule allows to classify types.

Fischbacher et al. (2001) - Idea

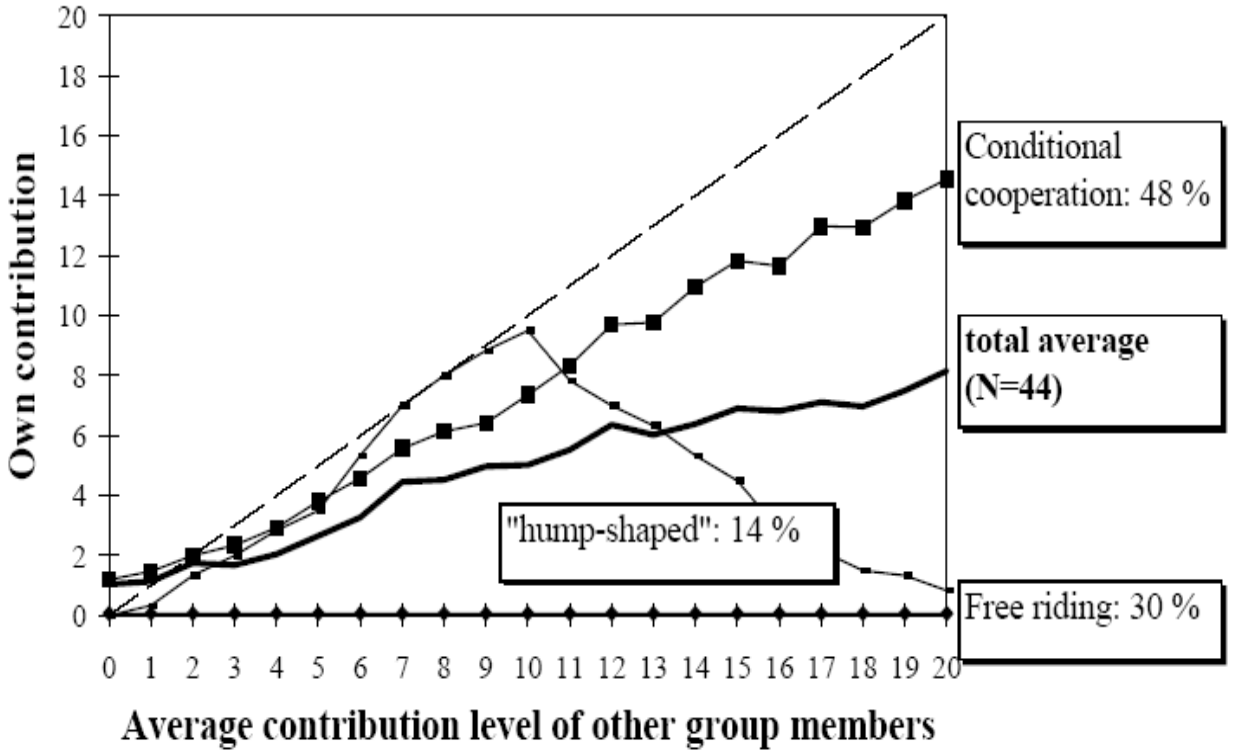


Fischbacher et al. (2001) - Types

- **Free riders:** They contribute nothing for any average group contribution .
- **Conditional cooperators:** They submit a contribution schedule that is monotonically increasing with the average group contribution.
- **Hump-shape contributors:** They submit a monotonically increasing contributions schedule up to an average group contribution of $x < 20$. Above x conditional contributions are monotonically decreasing.
- **Other contributors:** People who do not show a clear pattern in their contribution schedule.

Fischbacher et al. (2001) - Types

Average own contribution level for each average contribution level of other group members (Source: Fischbacher, Gächter & Fehr EL 2001)



Fischbacher and Gächter (2010)

- Can the observed pattern of conditional cooperation explain the unraveling of cooperation?
- Fischbacher and Gächter (2010) elicit types in an experiment (the P-experiment) and let players then play in a repeated public goods game (the C-experiment) to check whether types and beliefs predict behavior correctly.
- They show that contributions decline because, on average, people are “imperfect conditional cooperators” who match others’ contributions only partly.
- The presence of free-rider types is not necessary for this result; contributions also decline if everyone is an imperfect conditional cooperator.

Summary public goods experiments II

What influences contributions in linear public goods games?

Positive: marginal per capita return, communication, reward, punishment (monetary and non-monetary), positive framing, partner design, thresholds, endogenous association, leadership

Negative: heterogeneous endowments, experienced subjects, beliefs about others solicited, repetition, economic training

No significant effect: number of interaction periods, group size, gender, culture

And probably many more…

Summary public goods experiments III

The discussion about why people contribute and about some of the effects on contribution levels are not finally settled in economics.

The linear VCM is only one possible setting and probably not the most realistic.

Two overview articles:

Ledyard, J. O. (1995). Public goods. A survey of experimental research. In: Kagel and Roth (eds.). *Handbook of Experimental Economics*. Princeton University Press

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics* 6: 299-310.

Punishment

- Informal versus formal sanctions.
- How are sanctions usually implemented in PGGs?
- One institution that potentially fosters cooperation.
- Earlier assessments: Ostrom, Walker and Gardner (1992) and Yamagishi (1988).
- Fehr and Gächter (2000) introduce the concept to economics.

Fehr and Gächter (2000) - main research questions

- Will free riders be punished when punishment is costly and will punishment intensity depend on the difference in contributions?
- Can punishment increase the contributions of potential free riders?
- What are the effects of punishment on efficiency?

Fehr and Gächter (2000) - design

- 2x2 design: punishment/no punishment and partner/stranger
- 4-person groups
- Finite horizon (10 periods)
- Within-subject design (punish vs. no punish)

Fehr and Gächter (2000) - design

- Without punishment:

$$\pi_i^1 = y - g_i + a \sum_{j=1}^{n} g_j,$$

- With punishment:

$$\pi_i = \pi_i^1[1 - (1/10)P^i] - \sum_{j \neq i} c(p_i^j).$$

- $y = 20$; $a = 0.4$
- Theoretic solutions (under different assumptions)?

$$\partial \pi_i / \partial g_i = -1 + a < 0 \qquad 0 < a < 1 < na$$

Fehr and Gächter (2000) – results strangers

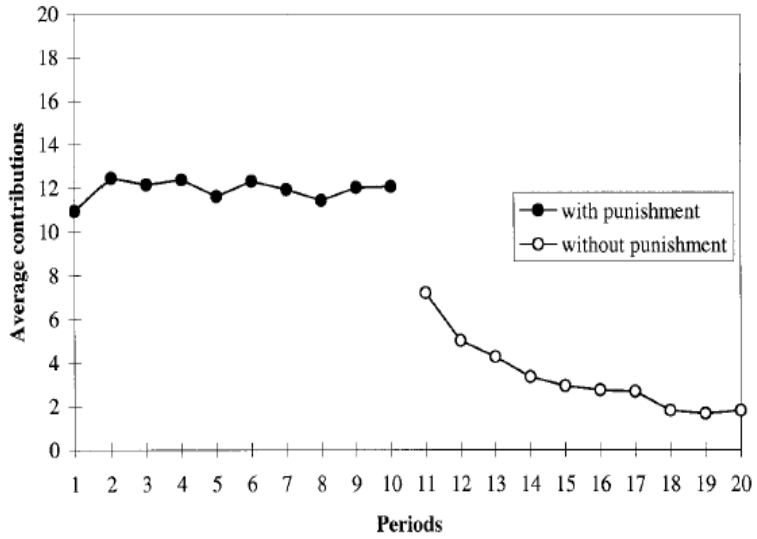


FIGURE 1A. AVERAGE CONTRIBUTIONS OVER TIME IN THE STRANGER-TREATMENT (SESSIONS 1 AND 2)

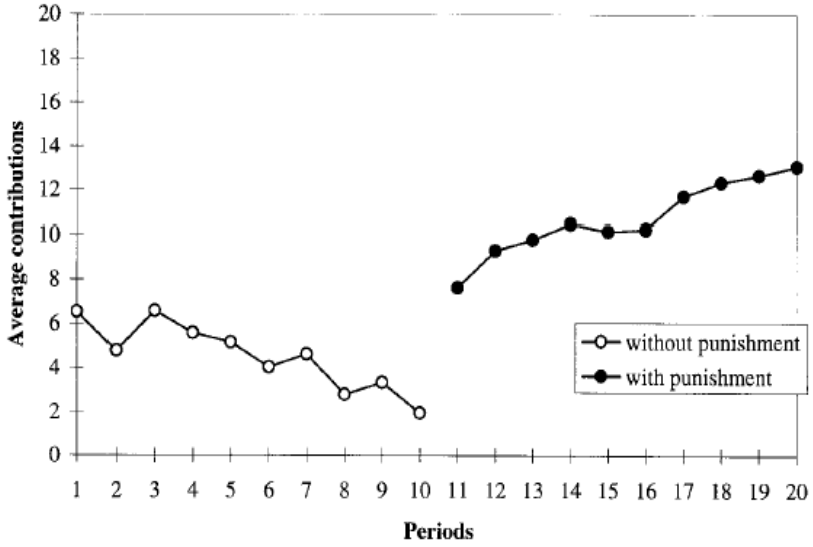


FIGURE 1B. AVERAGE CONTRIBUTIONS OVER TIME IN THE STRANGER-TREATMENT (SESSION 3)

Fehr and Gächter (2000) – results strangers

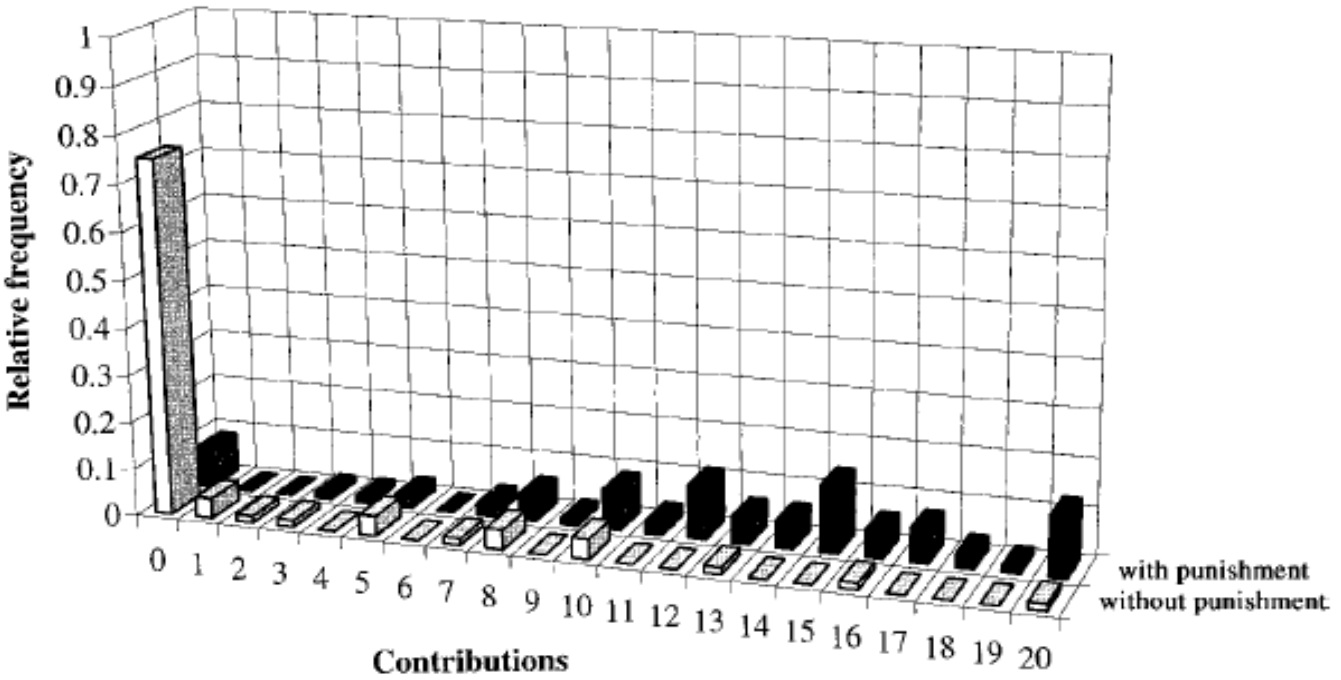


FIGURE 2. DISTRIBUTION OF CONTRIBUTIONS IN THE FINAL PERIODS OF THE STRANGER-TREATMENT WITH AND WITHOUT PUNISHMENT

Fehr and Gächter (2000) – results partners

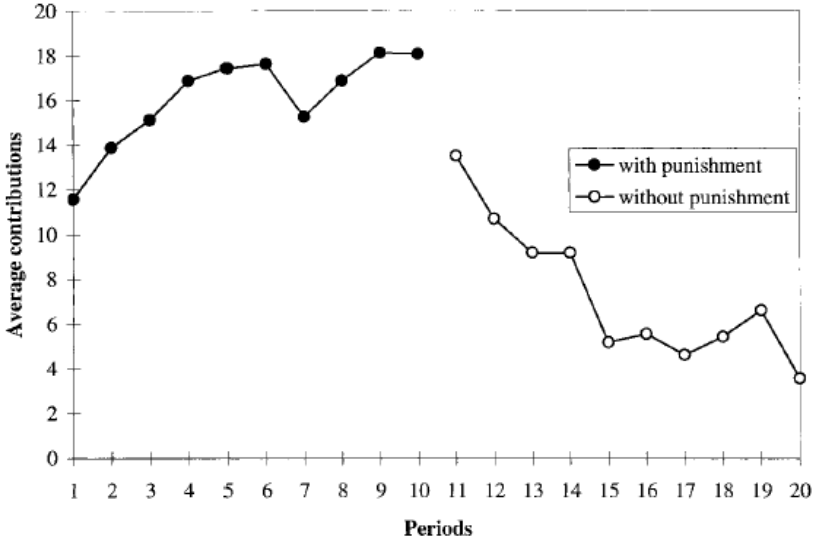


FIGURE 3A. AVERAGE CONTRIBUTIONS OVER TIME IN THE PARTNER-TREATMENT (SESSION 4)

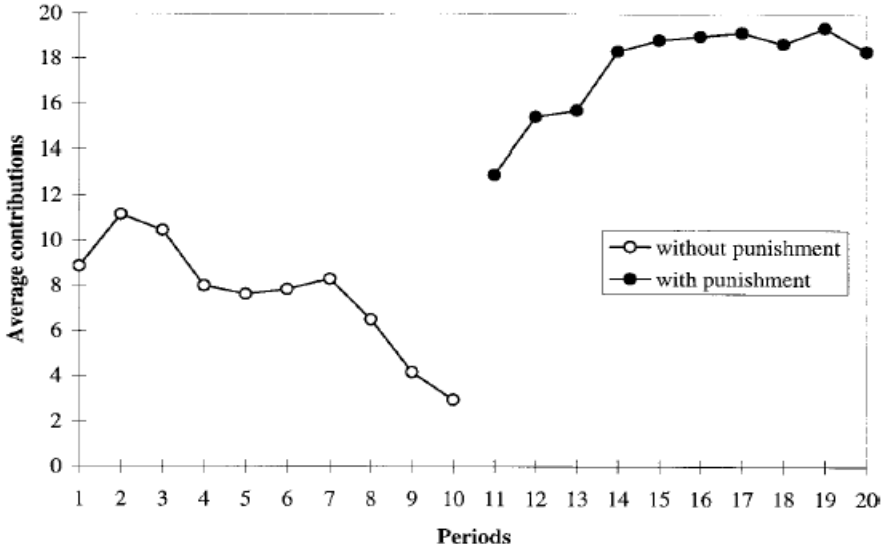


FIGURE 3B. AVERAGE CONTRIBUTIONS OVER TIME IN THE PARTNER-TREATMENT (SESSION 5)

Fehr and Gächter (2000) – results partners

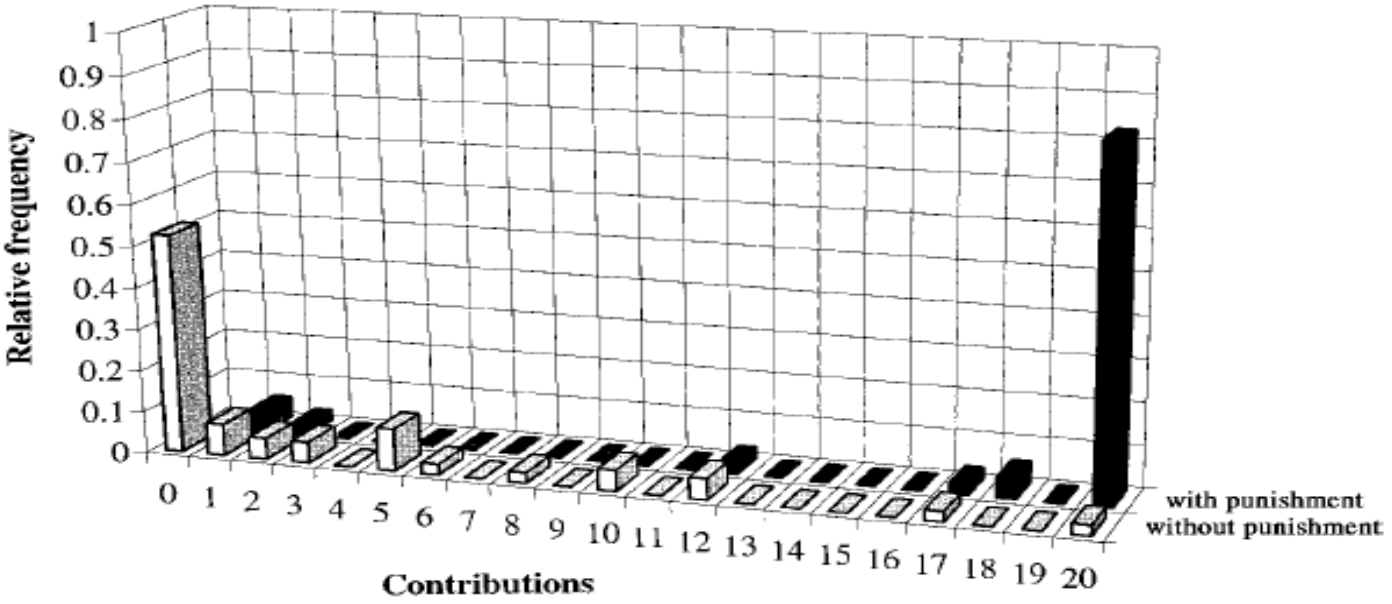


FIGURE 4. DISTRIBUTION OF CONTRIBUTIONS IN THE FINAL PERIODS OF THE PARTNER-TREATMENT WITH AND WITHOUT PUNISHMENT

Fehr and Gächter (2000) – results reaction functions

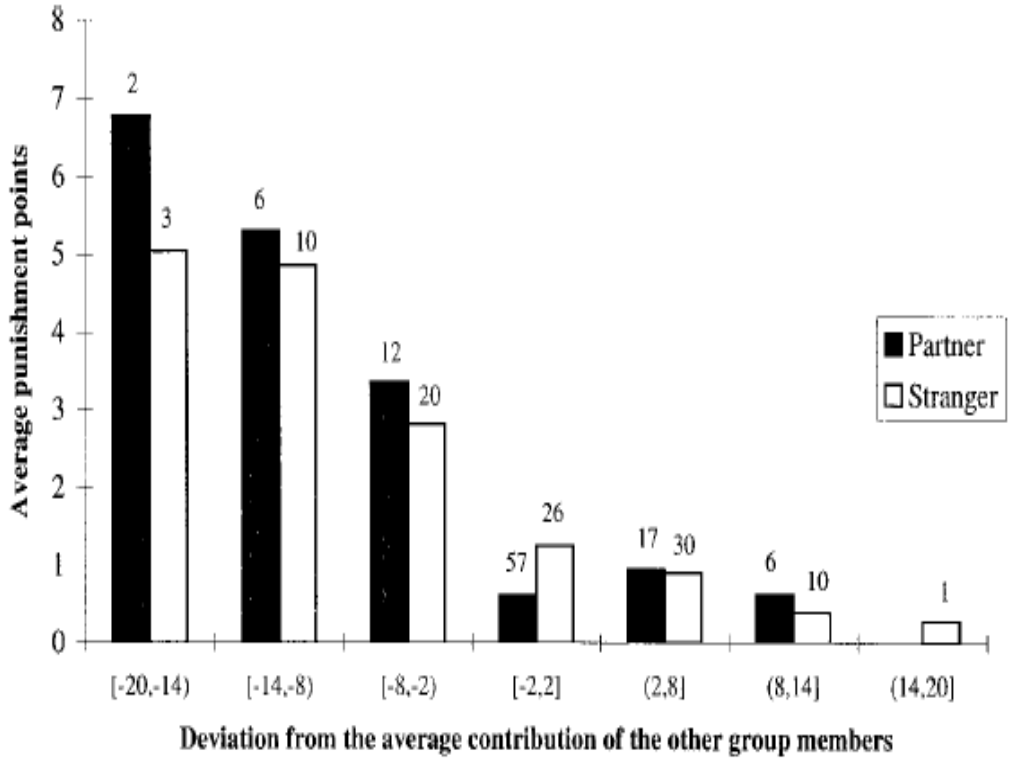


FIGURE 5. RECEIVED PUNISHMENT POINTS FOR DEVIATIONS FROM OTHERS' AVERAGE CONTRIBUTION

Fehr and Gächter (2000) – results reaction functions

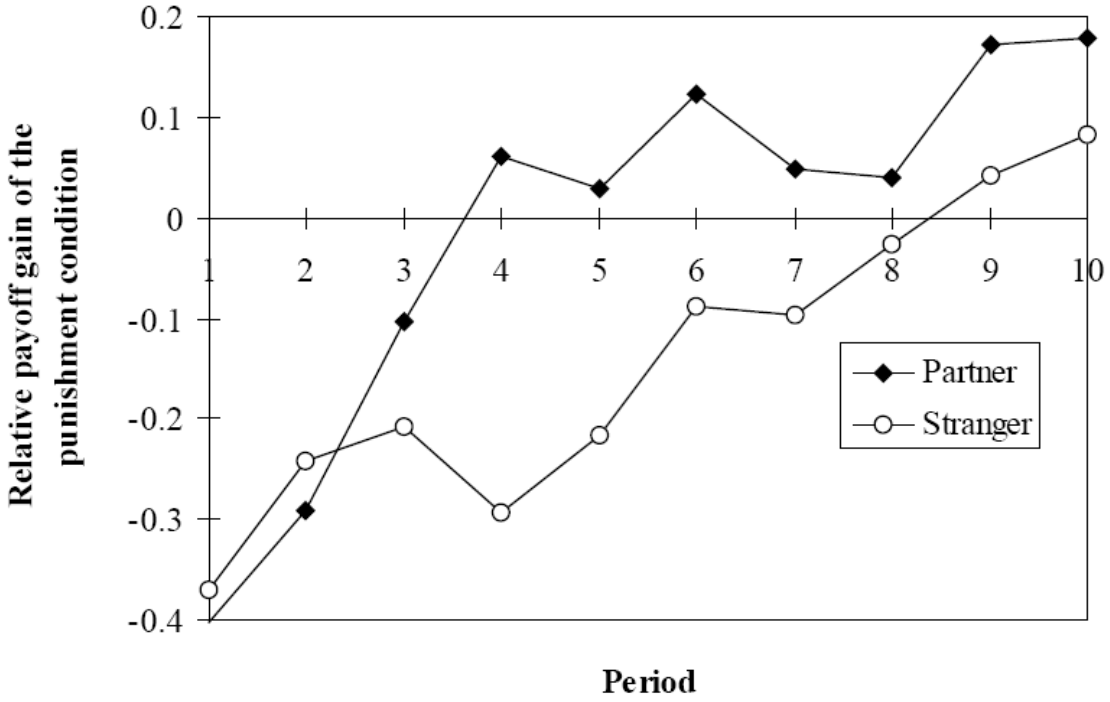
TABLE 5—DETERMINANTS OF GETTING PUNISHED: REGRESSION RESULTS

	Dependent variable: received punishment points	
Independent variables	Stranger-treatment	Partner-treatment
Constant	2.7363***	0.9881
	(0.0485)	(0.6797)
Others' average contribution	−0.0735***	−0.0108
	(0.0239)	(0.0457)
Absolute negative deviation	0.2428***	0.4168***
	(0.0325)	(0.0510)
Positive deviation	−0.0147	−0.0357
	(0.0264)	(0.0355)
	$N = 720$	$N = 400$
	$F[14, 705] = 39.0$***	$F[21, 378] = 41.3$***
	Adjusted $R^2 = 0.43$	Adjusted $R^2 = 0.68$
	DW = 1.96	DW = 1.89

Notes: Standard errors are in parentheses. * denotes significance at the 10-percent level, ** at the 5-percent level, and *** at the 1-percent level. To control for time and matching groups, the regression model also contains period dummies and dummies for matching groups (i.e., session dummies in the Stranger-treatment and dummies for each independent group in the Partner-treatment). Results are corrected for heteroskedasticity. Tobit estimations yield similar results.

Fehr and Gächter (2000) – results efficiency

Figure 6: Average payoff gain of the punishment relative to the no-punishment condition



Fehr and Gächter (2000) – summary and interpretation

- Punishment can sustain cooperation.
- What drives punishment? Emotions, social preferences, reciprocity,…?
- Unclear efficiency effects.
- Fehr and Gächter (2000) created “the” paradigm to study punishment in economics.

Sefton et al. (2007) - procedure

4 treatments:

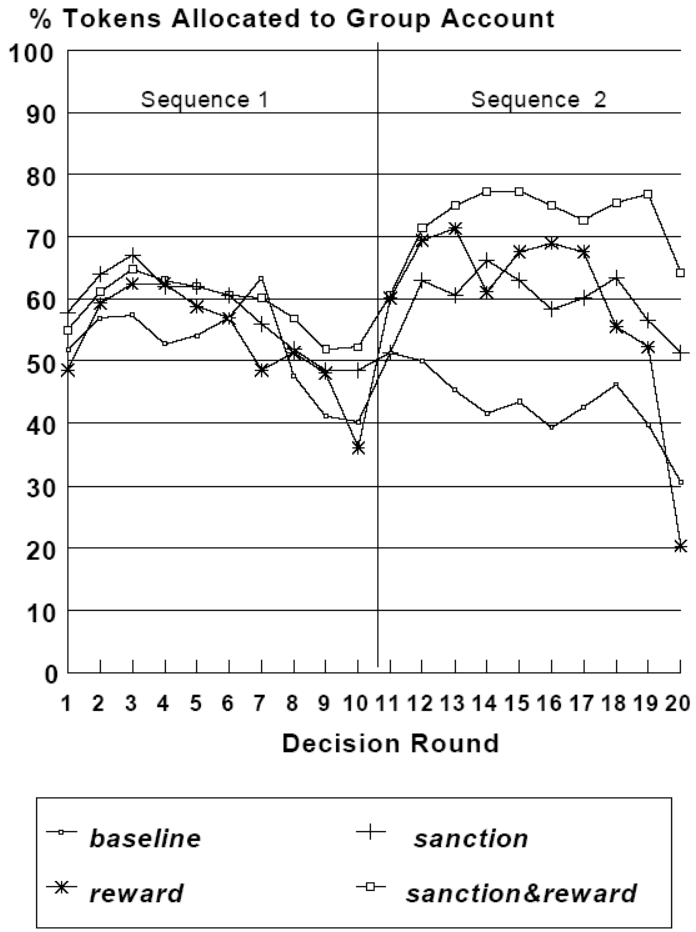
- VCM
- VCM plus punishment
- VCM plus reward
- VCM plus punishment and reward

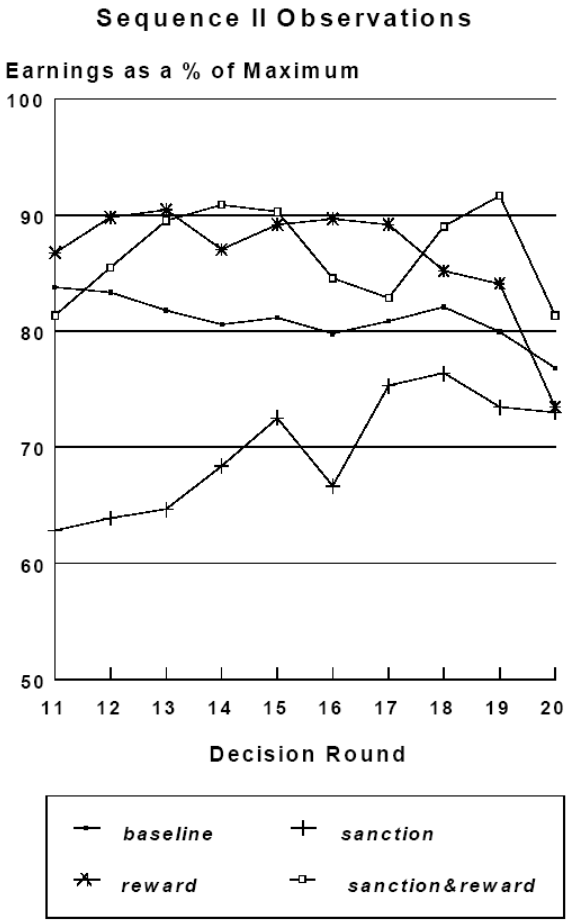
- Partner design (4-person groups)
- Each session consisted of two sequences (I. VCM, II. See above); 20 rounds.

Sefton et al. (2007) - procedure

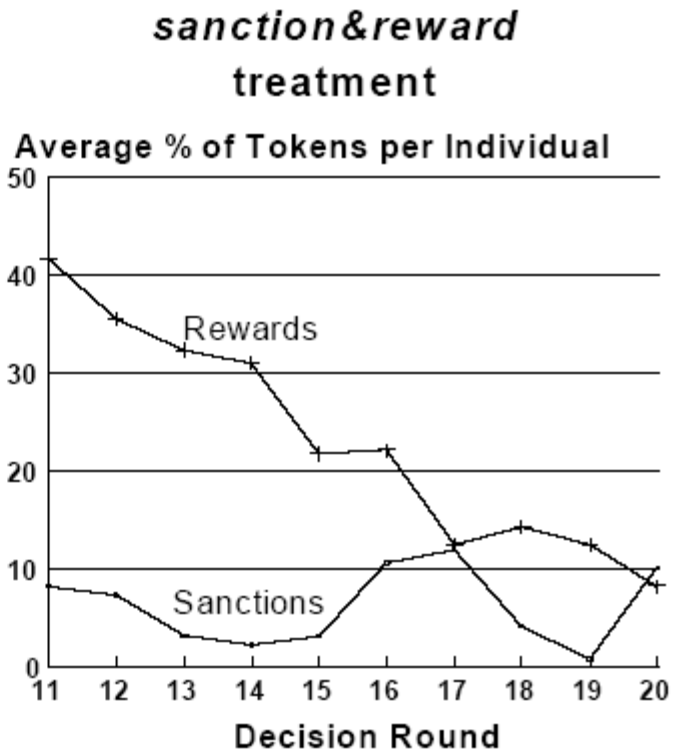
- Stage 1: 6 tokens per person to be allocated to the private account (10 cents per token) or the public account (5 cents per token)
- Stage 2 (when applicable): another 6 tokens to be kept (10 cents earnings per token) or to be used to punish/reward other group members (10 cents loss/10 cents gain).
- Introduction of punishment/reward effectiveness/costs (leverage).

Sefton et al. (2007) - results

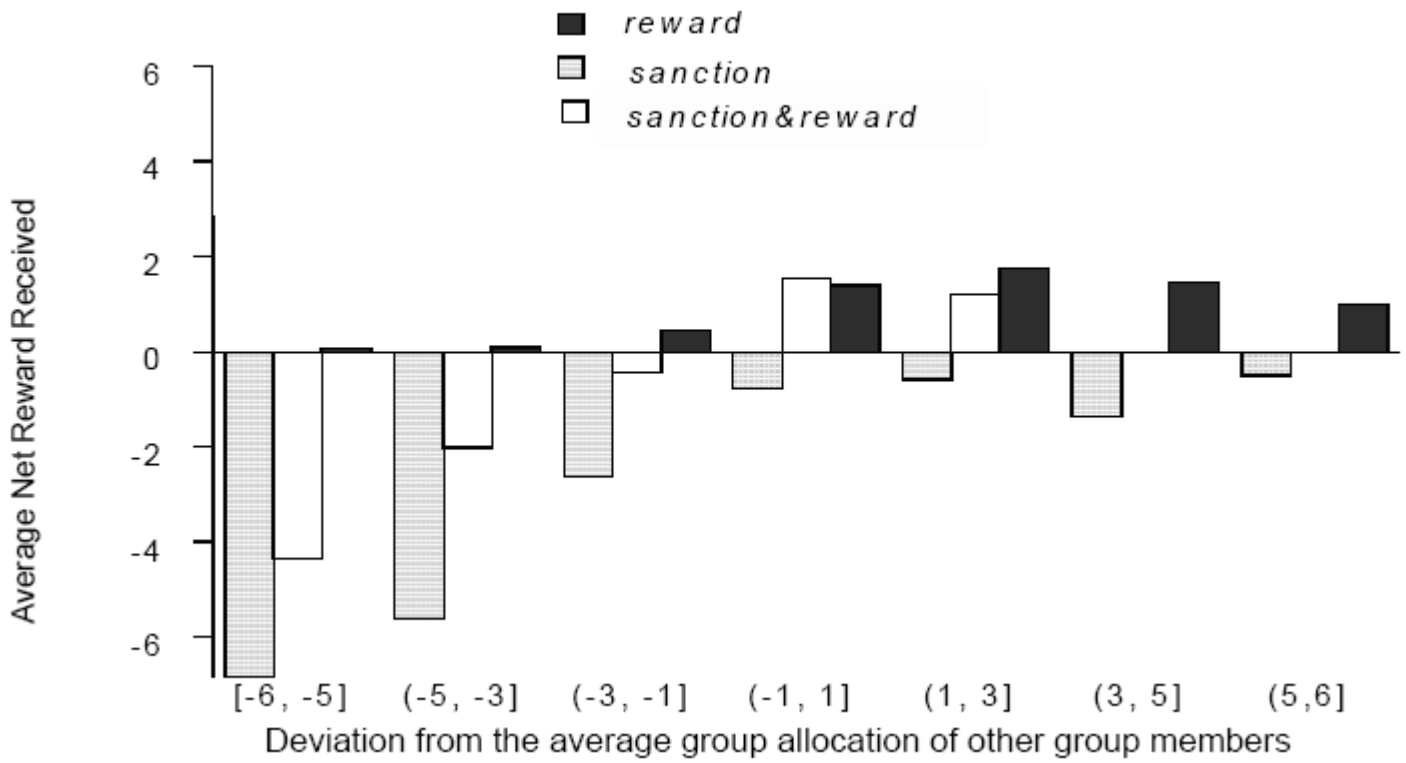




Sefton et al. (2007) - results



Sefton et al. (2007) - results



Sutter et al. (2010) – endogenous choice

- They examine the effects of endogenous institutional choice on behavior in social dilemma situations and address the following questions:
- Is there a marginal effect of endogenous institutional choice? I.e., does behavior depend upon endogenous or exogenous implementation of a rule?
- Which institutional rules are chosen endogenously (reward, punishment, neither)?
- Examples

Sutter et al. (2010) - design

- Groups with 4 subjects each (partner design)
- Endowment E = 20 tokens (= 0.8 €)
- Contribution to public good: $0 \leq c \leq 20$
- Payoff: $\pi_i = (20 - c_i) + 0.4 \Sigma c_j$
- 10 rounds
- Payoff with punishment/reward:

$$\pi_i = (20 - c_i) + 0.4\Sigma c_j \text{ +/- } L\Sigma p_{ki} - \Sigma p_{ik}$$

Sutter et al. (2010) - design

1. Control: Pure public goods game (no reward, no punishment) – only contribution phase (standard voluntary contribution mechanism - VCM)
2. Punishment: Additional punishment phase: subjects can punish each other (Yes or no)
3. Reward: Additional reward phase: subjects can reward each other (Yes or no)

Sutter et al. (2010) - design

1. Exogenous institutional choice: The institution is exogenously imposed (control/punishment/reward).
2. Endogenous institutional choice: Groups choose the institution endogenously by unanimous vote.

   a. Subjects choose whether to vote (one-time costs of 10 tokens).

   b. Voters can support any of the 3 institutions.

   c. If support is unanimous, institution is implemented.

   d. Otherwise, repeat steps b. and c.

Sutter et al. (2010) - design

- Leverage = 1 ("*L*=1"):

  Costs of punishment/reward: 1 token

  Costs/benefits of being punished/rewarded: 1 token

- Leverage = 3 ("*L*=3"):

  Costs of punishment/reward: 1 token

  Costs/benefits of being punished/rewarded: 3 tokens

- In the endogenous treatments, the leverage is fixed.

Sutter et al. (2010) - design

Choice of institution	Leverage	Control	Reward	Punishment
Exogenous	L = 1	N = 10	N = 10	N = 10
	L = 3		N = 10	N = 10
Endogenous	L = 1	N = 40 overall		
	L = 3	N = 20 overall		

Sutter et al. (2010) - predictions

Assuming common knowledge of selfishness and payoff maximization a sub-game perfect equilibrium is where all members contribute $c_{i,t} = 0$ in each round.

This result applies also to the treatments with punishment or reward because these actions are costly.

Prediction 1: Subjects contribute zero all the time.

Prediction 2: No subject ever participates in the vote on the institution in the endogenous treatment (because voting is costly).

Sutter et al. (2010) - predictions

- They consider the social preferences-model of Fehr and Schmidt (1999) to derive alternative predictions.
- If $x = (x_1, \ldots, x_n)$ denotes the vector of monetary payoffs of the n subjects of a group, Fehr and Schmidt (1999) define subject *i*'s utility U_i as follows:

$$U_i(x) = x_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_j - x_i, 0\} - \beta_i \frac{1}{n-1} \sum \max\{x_i - x_j, 0\}$$

  - It is assumed that $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$.

Sutter et al. (2010) - predictions

Contributions in the standard VCM:

If $\gamma + \beta_i < 1 \Rightarrow c_{i,t} = 0$

Let f denote the number of players with $\gamma + \beta_i < 1$

If $f/(n\text{-}1) > \gamma / 2 \Rightarrow c_{i,t} = 0$ for all i

If $f/(n-1) < (\gamma + \beta_i - 1)/(\alpha_i + \beta_i)$

$(\gamma + \beta_i > 1)$- players contribute $c_{i,t} \in [0, E]$

$(\gamma + \beta_i < 1)$- players contribute $c_{i,t} = 0$

Sutter et al. (2010) - predictions

Punishment with $(n' < n)$:

enforcer : $\gamma + \beta_i > 1$ and n' denotes the number of enforcers

Intuition:

For contributing $c_{i,t} = c > 0 \ \forall \ i$ to be an equilibrium it would be necessary that:

(1) contributing c is rational for selfish-players and

(2) that the threat of punishment is credible in case $c_{i,t} < c$

Sutter et al. (2010) - predictions

Punishment with $(n' < n)$:

(1) $c \leq \bar{c} = \frac{n'L}{(1-\gamma)}$

(2) Condition for punishment to be credible

$$-k - \frac{\alpha}{n-1}(n-n'-1)k - \frac{\alpha}{n-1}(\bar{c} - c + k - n'L) \geq$$

$$-\frac{\alpha}{n-1}(\bar{c} - c - (n'-1)L) - \frac{\beta}{n-1}(n'-1)k$$

$$\frac{L}{k} \geq (n-n') + \frac{1}{\alpha}[(n-1) - \beta(n'-1)]$$

$\Rightarrow$ Can only be fulfilled for *L* = 3 and $(n' > 1)$

Sutter et al. (2010) - predictions

Reward with $(n' < n)$:

enforcer : $\gamma + \beta_i > 1$ and n' denotes the number of enforcers

Intuition:

For contributing $c_{i,t} = c > 0 \ \forall \ i$ to be an equilibrium it would be necessary that:

(1) contributing c is rational for selfish-players and

(2) that the incentive to reward is credible in case $c_{i,t} = c$

Sutter et al. (2010) - predictions

Reward with $(n' < n)$:

(1) $c \leq \bar{c} = n'L/(1-\gamma)$

(2) Condition for reward to be optimal

$$(n'-1)L-(n-1)k-\frac{\alpha}{n-1}[L+(n-1)k](n-n') \geq$$

$$(n'-1)L-\frac{\beta}{n-1}[L-(n'-1)k](n'-1)$$

$$(n-1)k+\frac{1}{n-1}(L+(n-1)k)[\alpha(n-n')-\beta(n'-1)] \leq 0$$

$\Rightarrow$ Never fulfilled regardless of *L*

Sutter et al. (2010) - predictions

- Prediction 3: There is no difference between Control and Reward/Punishment with $L = 1$.

- Prediction 4: Subjects do not vote with $L = 1$.

Sutter et al. (2010) - predictions

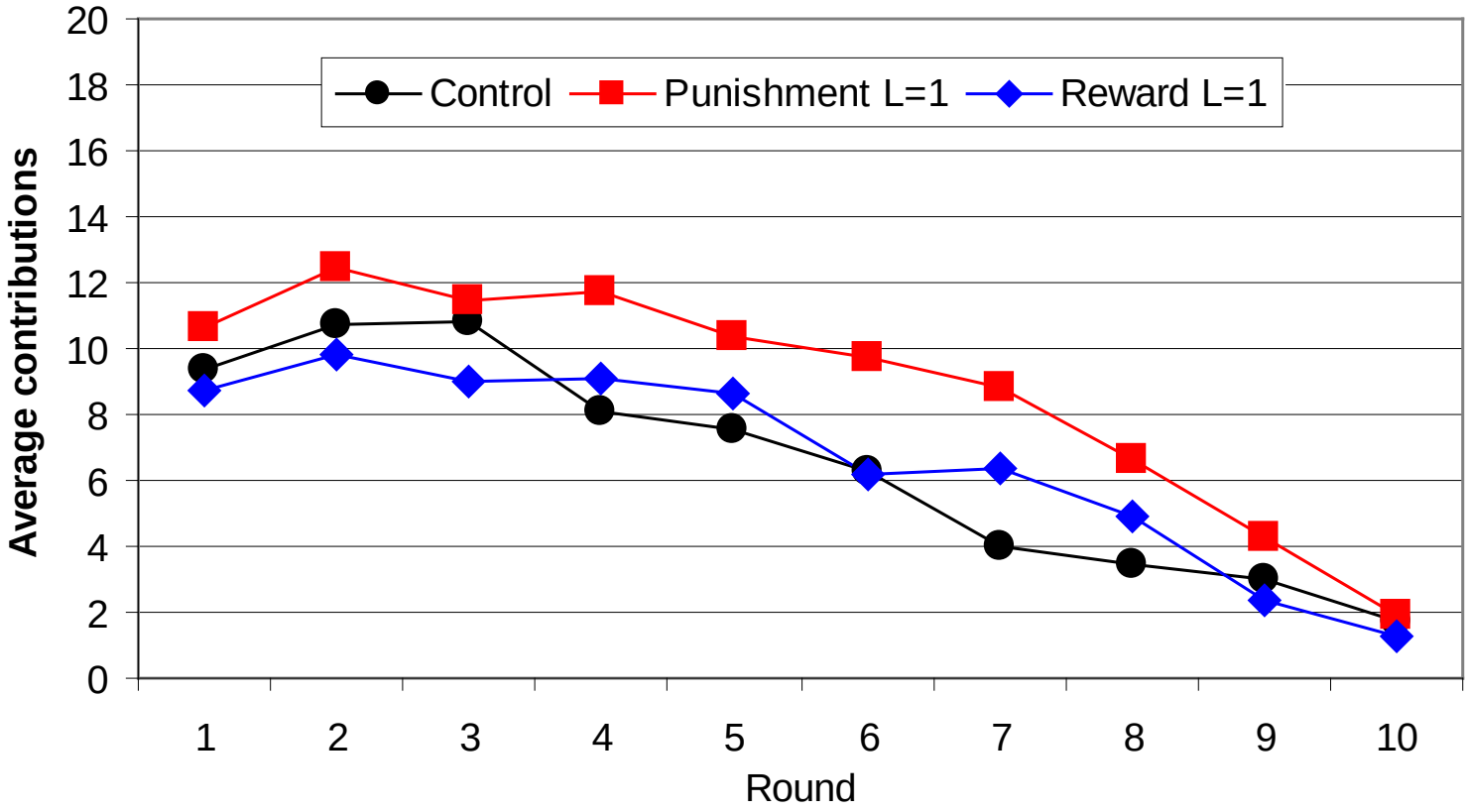
- Prediction 5: The contribution level $c^* = n'L / (1\text{-}0.4)$ can be enforced under punishment with $L = 3$, but not with reward.

  For $n' = 3$ we get $c^* = 15$.

- Prediction 6: Subjects should vote and should vote for punishment (for almost all possible parameter values of α_i and β_i (disregarding the coordination problem).

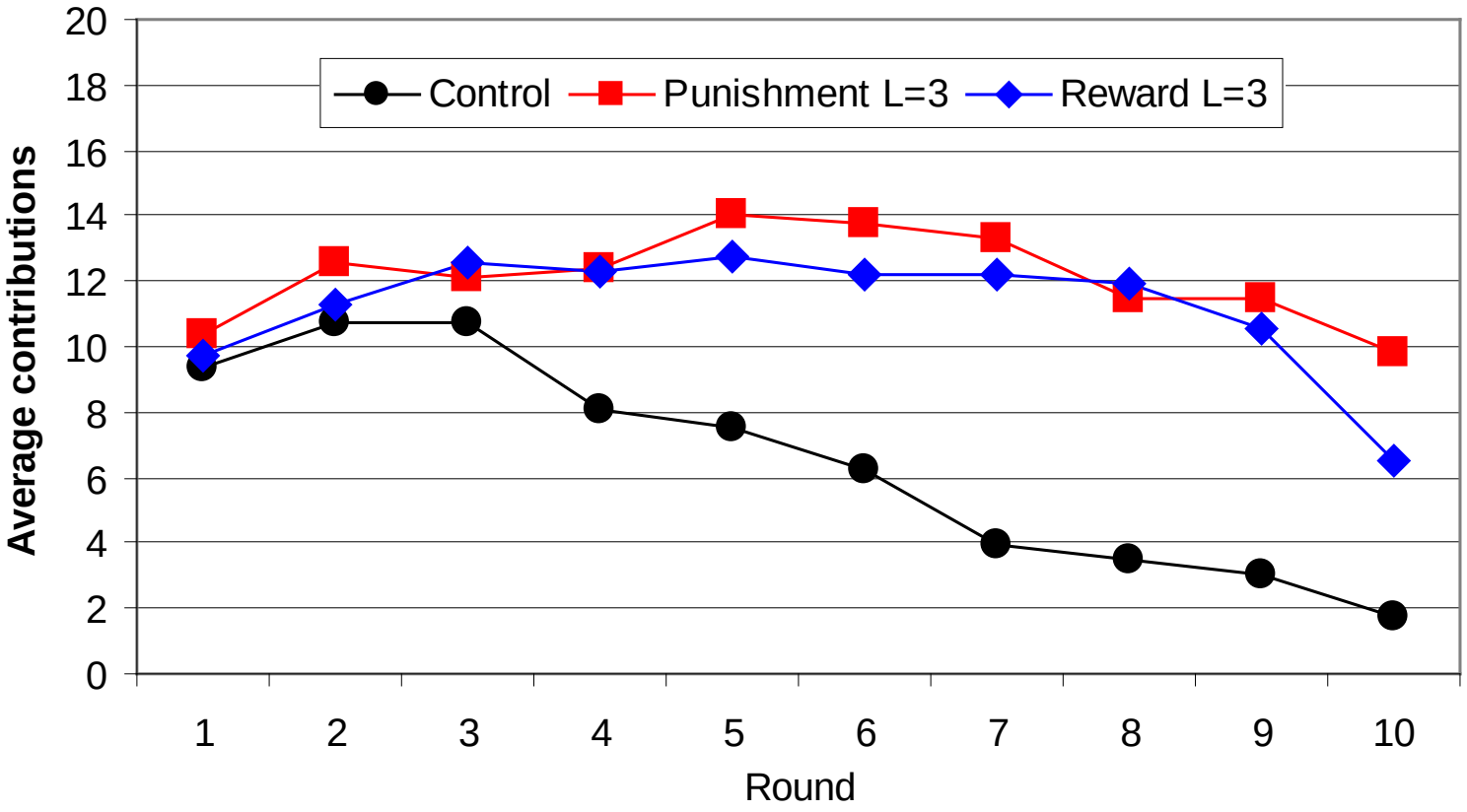
- Note: For $n' = n$: Every $c_i = c \in [0,E]$ is an equilibrium.

Sutter et al. (2010) – exo. results



No significant differences (see Predictions 1 and 3).

Sutter et al. (2010) – exo. results



Punishment and Reward significantly larger than Control (see Predictions 1 and 5).

Sutter et al. (2010) – exo. results

	Contributions	Profits	Rel. freq. of reward/punishm.
Control	6.50	23.9	-
Punishment *L*=1	8.79	24.3	0.15
Reward *L*=1	6.62	24.0	0.10
Punishment *L*=3	12.11	24.8	0.20
Reward *L*=3	11.20	29.4	0.44

Sutter et al. (2010) – results

Rel. frequency	*L* = 1	*L* = 3
Voters	0.44	0.60
Non-Voters	0.56	0.40
Participants (N)	160	80

Significantly higher voter turnout in *L* = 3

(see Predictions 2, 4 and 6).

Sutter et al. (2010) – results

Groups with institution	Control	Punishment	Reward
$L = 1$	25	5	10
$L = 3$	3	0	17

Institutional choice is far from random. Under $L = 1$ the majority of groups opt for Control, whereas under $L = 3$ the reward institution is predominant (see Prediction 6).

Sutter et al. (2010) – results

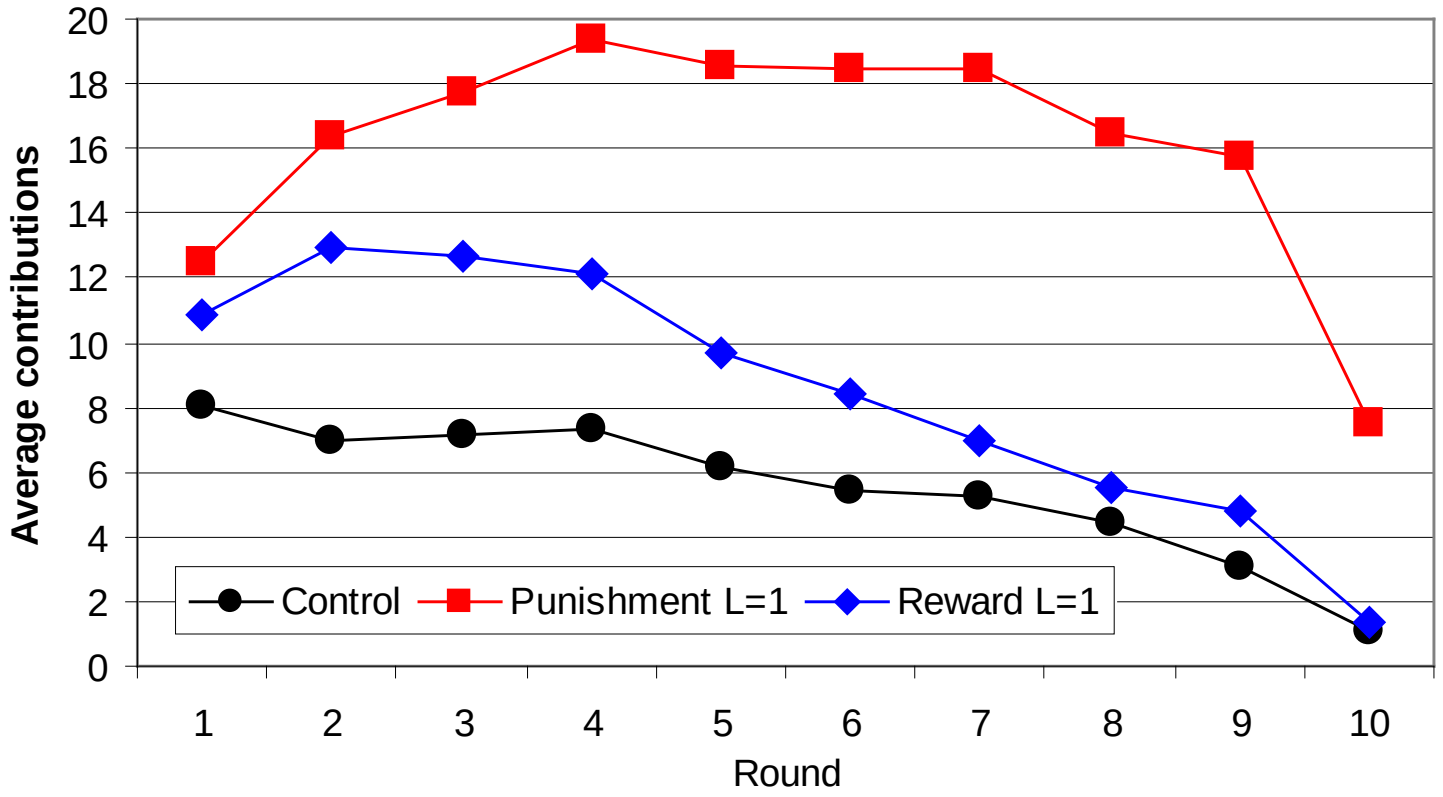
L = 1 (*N* = 40)

Voters		R 1	R 2	R 4	R 7	R 9	R 10	R 11	R 13	R 26
1	(15)	15								
2	(13)	9	1		1		1		1	
3	(6)	3	1					1		1
4	(3)		1	1		1				

L = 3 (*N* = 20)

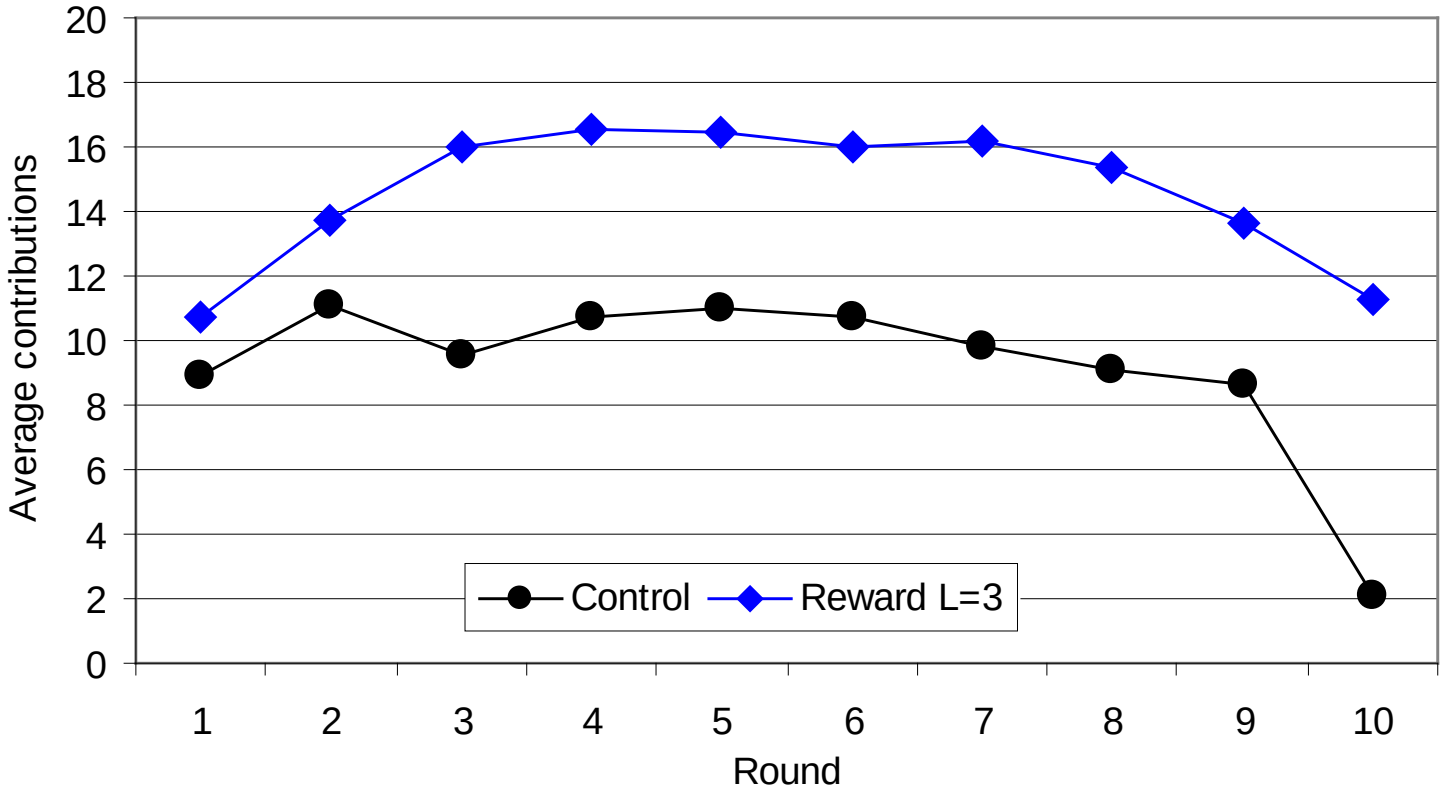
Voters		R 1	R 2	R 3	R 11	R 13	R 51	R 61	R 103
1	(5)	5							
2	(6)	5		1					
3	(5)	2	1			1		1	
4	(4)			1	1		1		1

Sutter et al. (2010) – endo. results



Punishment > Reward > Control (significantly).

Sutter et al. (2010) – endo. results



Reward > Control (significantly).

Sutter et al. (2010) – endo. results

	Contributions	Profits	Rel. freq. of re-ward/punishm.
Control "*L*1" (*N*=25)	5.48	23.3	-
Reward "*L*1" (*N*=10)	8.53	25.1	0.17
Punishment "*L*1" (*N*=5)	16.13	29.3	0.10
Control "*L*3" (*N*=3)	9.17	25.5	-
Reward "*L*3" (*N*=17)	14.59	32.1	0.55

Sutter et al. (2010) – results

$L = 1$	Control	Reward	Punishment
Exogenous	6.50	6.62	8.79
Endogenous	5.48	8.53	16.13
p (U-test)	n.s.	$p = 0.10$	$p < 0.01$

$L = 3$	Control	Reward	Punishment
Exogenous	6.50	11.20	12.11
Endogenous	9.17	14.59	-
p (U-test)	n.s.	$p = 0.09$	-

Endogenous choice of Reward or Punishment raises contributions.

Sutter et al. (2010) – results

TABLE 4

Determinants of contributions (censored tobit panel regressions)

Independent variables	(1) Exogenous treatments	(2) Endogenous treatments	(3) Exogenous plus endogenous treatments
Constant	13.81**	10.79**	10.50**
Period	−1.33**	−1.16	−1.24**
Punishment $\|L\| = 1$	2.43*	13.31**	7.41**
Reward $\|L\| = 1$	0.08	3.92**	2.88**
Punishment $\|L\| = 3$	5.51**	—	8.41**
Reward $\|L\| = 3$	5.35**	10.26**	8.70**
Voter	—	1.19	1.22**
Number of other voters in group	—	0.03	0.10
Endogenous treatment	—	—	1.71*
N	2000	2400	4400
Log likelihood	−4276.27	−4734.65	−9020.48

Note: *significant at 5% level; **significant at 1% level. Marginal effects shown.

Sutter et al. (2010) – results

TABLE 6

Punishing or rewarding behaviour (probit regression)

Independent variables	Dependent variable: decision to...		
	punish with $\|L\| = 1$	reward with $\|L\| = 1$	reward with $\|L\| = 3$
Constant	−0.901**	−1.692**	−2.148**
Voter (=1)	−0.173	0.217*	0.340**
Number of voters in group	0.238*	0.265**	−0.054
Group contribution	−0.017**	0.005**	0.038**
Difference between other members' and own contribution	−0.098**	0.035**	0.022**
N	600	1200	2040

Note: *significant at 5% level; **significant at 1% level.

Sutter et al. (2010) – results

TABLE 8

Contributions in the control experiment (censored tobit panel regression)

Independent variables	(1) Control experiment only	(2) Control experiment plus endogenous treatments with $\|L\| = 3$
Constant	4.81**	10.24**
Period	−0.38**	−0.25**
Punishment $\|L\| = 3$	1.89#	2.51**
Reward $\|L\| = 3$	3.04**	4.07**
Voter	1.66**	1.01*
Number of other voters in group	1.06**	0.55*
Vote implemented (1 = yes; 0 = non-preferred institution implemented by chance move)	0.76	—
Control experiment (1 = yes)	—	−5.32**
N	1600	2400
Log likelihood	−3791.77	−5379.82

Note: **Significant at 1% level; *significant at 5% level; #significant at 6% level. Marginal effects shown.

Sutter et al. (2010) – conclusion

- Reward and punishment enhance cooperation in public goods games, but only when the leverage of these institutions is high or when they are chosen endogenously. Hence, endogenous choice is a substitute for higher leverage of reward or punishment – and can, therefore, avoid the inefficiencies associated with punishment (cf. Nikiforakis and Normann, 2005).

Sutter et al. (2010) – conclusion

- If given the choice, subjects opt for reward when it has a high leverage (L = 3), but mostly prefer the standard public goods game when the leverage of reward or punishment is low (L = 1). Punishment is an unattractive institution (cf. Botelho et al., 2005), but very effective, if chosen endogenously.
- Endogenous choice of punishment might make the threat of punishing free-riders more credible.

Sutter et al. (2010) – conclusion

- In sum, endogenous institutional choice adds a “democratic participation rights”-premium to contributions under reward or punishment.
- This cooperation-increasing effect of endogenous choice is independent of the number of voters in a group.
- A subject’s decision to vote on the institution is a good indicator of her (higher) level of cooperation (in the beginning).
- Theoretical implications.