BACKGROUND ON PUBLIC GOODS

- A good is called *non-rival* of consumption of it by one individual does not diminish the amount available for consumption by other individuals.
  - Example: satellite TV broadcast.
- A good is called *non-excludable* if no individual can be prevented from consuming it (or can only be prevented at an unreasonably high cost).
  - Example: national defense.
- *Private goods* are both rival and excludable.
  - Example: food.
- *Pure public goods* are both non-rival and non-excludable.
  - Example: outcomes of basic scientific research.
- *Common-pool resources* are rival but non-excludable.
  - Example: public park.
- *Club goods* are non-rival but excludable.
  - Example: coded satellite TV broadcast.
- Table classification:

Rivalry/Excludability	**Non-excludable**	**Excludable**
Non-rival	Pure public good	Club good
Rival	Common-pool resource	Private good

- Private sector is efficient, via markets, in producing and delivering private goods. But not public goods. Non-excludability implies that anyone, even the ones who did not pay for it, can consume a public good. Because of that, everyone has an incentive to do exactly that, i.e., to *free-ride.*
- As a result, if the provision public goods is left up to the private sector, there will generally be *underprovision*: anyone willing to pay for a bit of the public good ignores the positive consumption externality such act infers on other people and hence too little of the good is purchased.
- In fact, this has been documented experimentally via means of **Voluntary Contribution Mechanism (VCM)**.
- In reality, public goods are typically provided by various levels of government that also levy (often distortive) taxes to finance these goods.
- The follow-up literature on VCM tries to identify institutional designs that would lead to private provision of public goods. We will in turn consider two of those:
  - **VCM with Punishments for Non-Cooperation**
  - **Provision-Point Mechanism**

VOLUNTARY CONTRIBUTIONS MECHANISM (VCM)

- Consider the following game of n players: each player is endowed with a budget of $y > 0$. Each player i needs to split this budget between a contribution to the public account g_i and his private good consumption $y - g_i$. The sum $\sum_i g_i$ is a metaphor for the amount of the produced public good. Each player then receives this sum multiplied by a factor of $a \in (1/n, 1)$. This factor, also called a **marginal per capita return (MPCR)**, is a metaphor for the marginal utility (in terms of private consumption) of the public good. Hence the ultimate payoff of player i is given by

$$x_i(g_1, ..., g_n) = y - g_i + a \sum_{j=1}^{n} g_j.$$

Because $a < 1$, the marginal cost of contributions to public account (in terms of private consumption), in this game it is a dominant strategy to contribute nothing, and this is the unique prediction of game theory. On the other hand, since $na > 1$, it is socially optimal if everyone contributes everything. Given this tension between private and social interest, this game is an n-player Prisoner's Dilemma.

Marwell and Ames (1981)

- This is an early study of this kind. The authors consider various variations on this game and this is what they find:
- We observe that:

Table 2

Summary of results: Experiments 1–11.

Experiment	Mean % of resources invested
1. Basic experiment	42%
2. Skewed resources and/or interest	53%
Experiments 1 and 2, combined	51%
3. Provision point	51%
4. Small groups with provision point (except those with sufficient interest to provide the good themselves)	60%
5. Experienced subjects	47%
6. High stakes	
Experienced interviewers	35%
All interviews	28%
7. Feedback, no changing initial investment	46%
8. Feedback, could change investment in individual account	50%
9. Feedback, could change investment in individual account — college students	49%
10. Manipulated feedback	
Low	43%
Medium	50%
High	44%
11. Non-divisibility	
Divisible (control)	43%
Non-divisible	84%
12. Economics graduate students	20%

1. The contribution rate is around 40% to 50% of available resources, contradicting the game-theoretic prediction.
2. The only exception is when the experiment is done with a group of doctoral students in economics, who contribute only around 20% of available resources.

- Note that the paper is titled "Economists Free Ride, Does Anyone Else?"

Isaac and Walker (1988a)

- The authors are motivated by the conjecture that larger groups have a harder time contributing toward the public good. They argue that this comparative static may be based on thinking of a fixed factor na that multiplies the pot of public good contributions. As a result, if n increases, a, or MPCR must fall, which may lead to lower contributions. However, from the real world point of view, a more relevant exercise is to see what happens when a is kept constant while n increases.

- In particular, the authors experimentally study comparative statics of the VCM game with respect to both n and a, *ceteris paribus*. They use a within-subject design, running two series of 10 periods with $a = 0.3$ and $a = 0.75$ (with changing the order in half of sessions to control for order effects).

- Here are the details of the design (note that the multiplier in the third column is equal to na and represents a factor by which the pot of public good contributions is multiplied before being equally split among all the players):

TABLE I
EXPERIMENT PARAMETERS

Experiment type	Group size	Group payoff function	MPCR	Individual tokens per period (Z_i)	Number of experiments
4*L*	4	$1.2(\Sigma m_i)$¢	0.30	62	6
4*H*	4	$3.0(\Sigma m_i)$¢	0.75	25	6
10*L*	10	$3.0(\Sigma m_i)$¢	0.30	25	6
10*H*	10	$7.5(\Sigma m_i)$¢	0.75	10	6

- The authors focus on identifying **strong free-riders**, i.e., subjects who contribute less than one third of their endowment. They also claim that their qualitative conclusions are insensitive to the particular threshold (if lower) that defines the category of strong free-riders.

- Results for all 10 periods:

- We observe that:

  1. Holding group size constant, lowering the MPCR from 0.75 to 0.3 significantly increases the incidence of free-riding behavior.
  2. Holding MPCR constant, there are weak, if any, effects of group size (4 vs. 10 players) on free-riding behavior.
  3. Similar conclusions obtain for the last-period effects of group size and MPCR:
  4. The level of free-riding tends to increase over time, which is consistent with results on Prisoner's Dilemma games we talked about earlier.

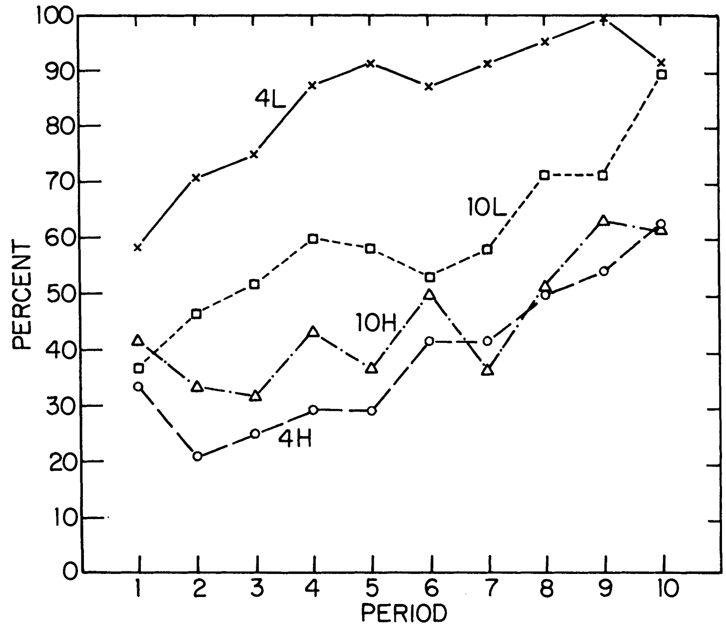

FIGURE I
Mean Percent of Individuals Acting as Strong Free Riders

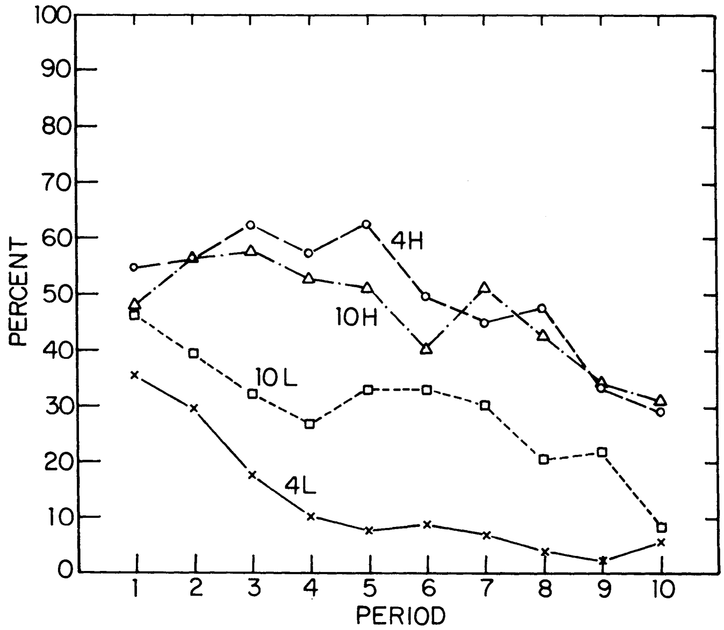

FIGURE II
Mean Percent of Tokens Contributed to the Public Good

TABLE II
END-PERIOD RESULTS

Experiment type	Replication	Percent of tokens contributed to the public good	Number of strong free-riders	Number of persons contributing zero
$4L$	1	0.0%	4	4
	2	0.0%	4	4
	3	0.0%	4	4
$4H$	1	25.0%	2	2
	2	52.0%	2	1
	3	10.0%	4	2
$10L$	1	0.4%	10	9
	2	11.6%	8	8
	3	10.0%	9	6
$10H$	1	25.0%	7	6
	2	21.0%	7	7
	3	25.0%	7	6

Isaac, Walker and Williams (1994)

- These authors extend Isaac and Walker (1988) but considering four different group sizes: $4, 10, 40$ and 100.
- Experimental design (MS = multiple session; SS = single session; XC = extra credit)
- Results:

R.M. Isaac et al., Group size and the voluntary provision of public goods 13

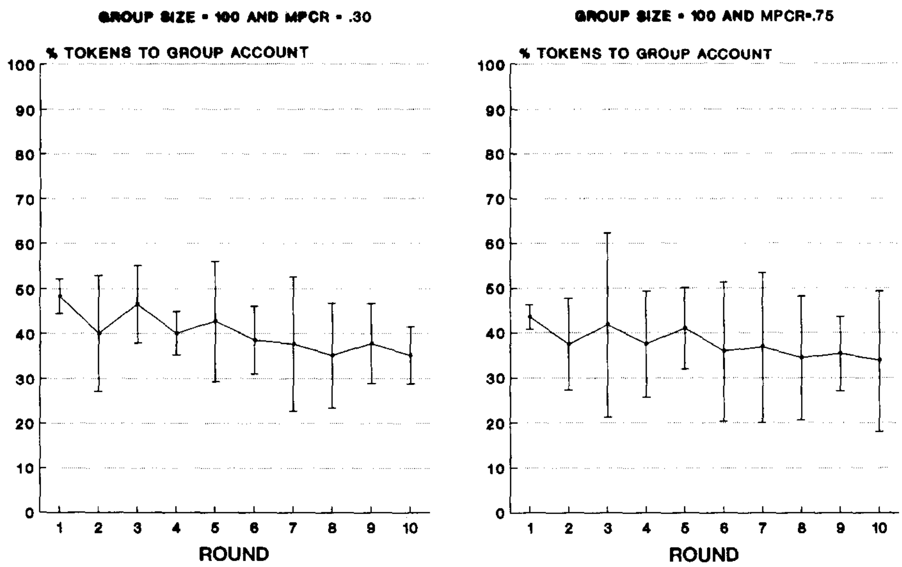

Fig. 5. 90% confidence bands: group size = 100.

- We observe that:
  1. For group size of 40 and 100, variation in MPCR seems not to affect the results.
  2. For MPCR=0.30, large groups (40,100) contribute more that small groups (4,10). For MPCR=0.75, there are no significant differences in the contribution rate across groups of different sizes.

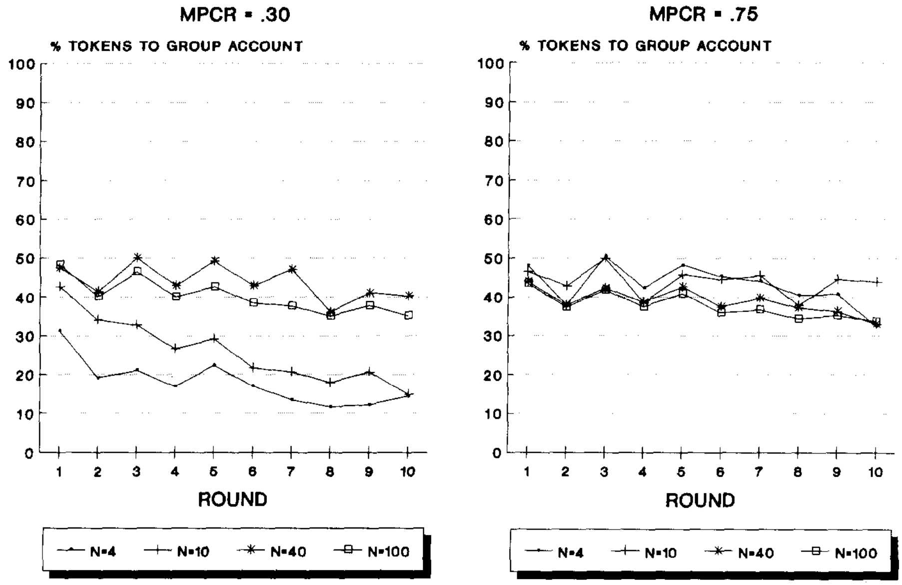

Fig. 6. Group size comparison for high (0.75) and low (0.30) *MPCR* cells.

Goeree, Holt and Laury (2002)

- This paper looks at the comparative statics of VCM with respect to MPCR. It argues, though, that a higher MPCR both increases the value of own contribution to the others and reduces the private marginal cost of contribution. The authors distinguish between these two effects and identify the impact of each of them separately.

- Analytically, using the same setup as before, the ultimate payoff of player i is given by

$$x_i(g_1, ..., g_n) = y - g_i + a_i g_i + a_{-i} \sum_{j \neq i} g_j.$$

  Here, a_i is the MPCR, or *internal return*, to own contribution, whereas a_{-i} is the *external return* to the contribution of others.

- Procedure: 10 rounds, strategy method, random rematching, no feedback until the end of the experiment $n = 2$ or $n = 4$, $y = 25$ tokens, a token kept in a private account yields 5 cents. Internal and external rates of return:

Table 1
Summary of treatments

	Treatment									
	1	2	3	4	5	6	7	8	9	10
Group size	4	2	4	4	2	4	2	2	4	2
Internal return	4	4	4	2	4	4	2	4	2	4
External return	2	4	6	2	6	4	6	2	6	12
Mean contribution	10.7	12.4	14.3	4.9	11.7	10.6	7.7	6.7	10.5	14.5
Median contribution	10	14	17	5	14	11	7	5	10	16.5

- Note that the internal rate of return always falls short of 5, so it is always a dominant strategy to contribute 0. On the other hand, the social rate of return is always more than 5, so it is always Pareto efficient to contribute everything.

- Results:

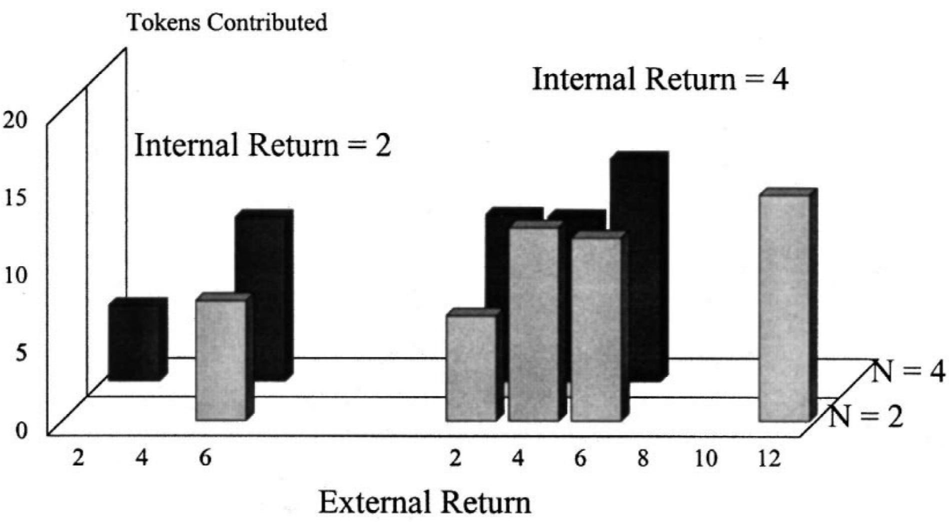

Fig. 1. Average contributions by treatment (number of tokens contributed).

- We observe that:

    1. *Ceteris paribus*, an increase in the internal rate of return from 2 (40%) to 4 (80%) has a strong positive impact on contributions.
    2. *Ceteris paribus*, contributions increase with increases in the external rate of return.
    3. *Ceteris paribus*, contributions increase with increases in the group size from $n = 2$ to $n = 4$.

- Implication: contributions do respond to the aggregate benefit generated by the contribution even though the action is always privately costly. This is consistent with subjects being altruistic, but not consistent with altruism being of the "warm-glove" type (utility purely from the act of giving rather than from what the others receive).

PROVISION-POINT MECHANISM

- This mechanism is suitable for provision of discrete public goods, but can also be used in a continuous public good environment.
- The idea of the provision-point mechanism is to modify the VCM by *ex ante* establishing a target level of contributions and then only providing the public good if this aggregate level of contributions is reached. If not, the contributions are simply refunded and no public good is provided. If more than the threshold amount is collected, several different things may be done:

  1. more public good may be provided,
  2. money may be refunded to contributors in equal amounts.

- Many fundraising campaigns rely on this kind of mechanism.
- Within the environment of the VCM, suppose that the target level is set at $\tau \in (0, ny]$. Then there can be several types on Nash equilibria on the modified contributions game:

  1. Any set of contributions that sum up to exactly τ.
  2. Any set of contributions that sum up to less than τ with the property that

     $$\tau - \sum_j g_j > y - g_i$$

     for all players i. That is, nobody can afford to top the existing contributions so as to push the total over the threshold.

- Note that if $\tau = ny$, then it is a weakly dominant strategy for all players to contribute y.

Dawes, Orbell, Simmons, Van de Kragt (1986)

- The authors consider a scenario with a discrete public good, all or nothing contributions and a flat bonus if the public good is provided. Although the paper is framed differently, the interesting aspect is that it compares the standard provision point mechanism (with *money-back guarantee* if the public good is not provided) with an augmented mechanism that forces contributions by non-contributors in case the public good is provided (referred to as *enforced contributions*). Note: the authors also consider a "standard dilemma" scenario in which the contribution money is burned in case the threshold is not reached.
- In this implementation, $y = 5$, $n = 7$, the contribution choice is discrete (all or nothing) and each subject is paid a bonus of 10 if the public good is provided.
- Payoff table of the "standard dilemma":

	Threshold reached	Threshold not reached
Contribute	10	0
Do not contribute	15	5

- With the money-back guarantee, it is

	Threshold reached	Threshold not reached
Contribute	10	5
Do not contribute	15	5

  Hence the part of strict domination of "Do not contribute" is removed if the threshold is not reached.

- With enforced contributions, it is

	Threshold reached	Threshold not reached
Contribute	10	0
Do not contribute	10	5

  Hence the part of strict domination of "Do not contribute" is removed if the threshold is reached.

- Results for the threshold of 3 contributions:

Table 1. Experiment 1: Percentage Contributing, Number of Contributors, and the Analysis of Variance and Scheffé Test Results (3 or 7 required)

	Standard Dilemma (1)	Money-Back Guarantee (2)	Enforced Contribution (3)
Percentage contributing	51	61	86
Number of contributors in each group	1, 2, 2, 3, 3, 3, 4, 5, 6, 7	3, 3, 4, 4, 5, 5, 6	4, 5, 5, 6, 6, 6, 7, 7, 7, 7

- We observe that enforced contributions lead to a significant improvement in contributions in comparison to money-back guarantee.

- Results for the threshold of 5 contributions:

Table 2. Experiment 2: Percentage Contributing, Number of Contributors, and the Analysis of Variance and Scheffé Test Results (5 of 7 required)

	Standard Dilemma (1)	Money-Back Guarantee (2)	Enforced Contribution (3)
Percentage contributing	64	65	93
Number of contributors in each group	3, 4, 4, 4, 4, 4, 5, 5, 6, 6	3, 4, 4, 5, 5, 5, 6	5, 6, 6, 6, 7, 7, 7, 7, 7, 7

- We observe that enforced contributions lead to a significant improvement in contributions in comparison to money-back guarantee.

Bagnoli and McKee (1991)

- This is an experimental implementation of Bagnoli and Lipman (1989). In this model, the public good is discrete and each subject has a private valuation for it. Contributions are continuous.

- Setup 1: $n = 5$, the cost of the public good is 12.5 and the sum of individual valuations is 25 (7 sessions, groups 11-17).
- Setup 2: $n = 10$, the cost of the public good is 25 and the sum of individual valuations is 50 (2 sessions, groups 20 and 21).
- 14 rounds, fixed groups.
- Results:

TABLE I
Total Contributions by Group—in Tokens

	Group Number								
Period	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**20**	**21**
1	20.0	10.5	15.0	12.5	17.0	24.0	18.0	38.0	29.5
2	14.5	13.0	11.0	12.5	15.2	16.5	14.1	28.5	25.5
3	12.0	12.5	14.5	12.5	11.5	12.0	13.2	23.3	25.0
4	13.0	12.0	13.5	12.5	10.0	12.0	12.5	17.2	24.0
5	12.5	12.5	11.0	12.5	13.5	15.0	12.5	23.5	19.5
6	12.0	10.0	12.5	12.5	12.8	14.0	12.5	25.5	23.5
7	12.5	13.0	12.5	12.5	12.8	13.5	12.5	25.5	24.5
8	12.5	12.5	12.0	12.5	12.5	13.0	12.5	26.5	26.5
9	12.5	13.0	12.5	12.5	12.5	13.5	12.5	24.0	25.0
10	12.5	12.3	12.5	12.5	12.5	13.0	12.5	25.2	26.0
11	12.5	13.0	12.5	12.5	12.7	13.0	12.5	24.25	25.0
12	12.5	12.0	12.5	12.5	12.5	13.0	12.5	25.0	25.0
13	12.5	12.5	12.5	12.5	12.5	13.0	12.5	25.0	25.0
14	12.5	13.0	12.5	12.5	12.5	12.5	12.5	25.0	28.5

- We observe that:
  1. With $n = 5$, over all rounds, public good is provided in 85 out of 98 cases and it is provided without any wasteful contributions in 53 out of 98 cases. In the last 5 rounds, public good is provided in 33 out of 35 cases and it is provided without any wasteful contributions in 26 out of 35 cases.
  2. With $n = 10$, over all rounds, public good is provided in 19 out of 28 cases and it is provided without any wasteful contributions in 8 out of 28 cases. In the last 5 rounds, public good is provided in 9 out of 10 cases and it is provided without any wasteful contributions in 6 out of 10 cases.
  3. Over all rounds, welfare levels are statistically significantly higher with $n = 5$ than with $n = 10$. However, this difference disappears in the last 5 rounds.
- Conclusion: subjects are capable of achieving the efficient provision of the public good via the provision-point mechanism. The only difference with respect to the group size is that larger groups take longer to learn and to coordinate on an efficient equilibrium.

COOPERATION-ENFORCING INSTITUTIONS: PUNISHMENTS FOR NON-COOPERATION

- The free-riding problem that obtains in experimental implementations of the linear public goods game with VCM stands in contrast to the casual observation that societies are often quite successful in achieving a high level of cooperation in various social dilemma situations, provision of public goods being probably the most important example. A crucial observation here, though, is that contributions are often enforced by a threat of and, sometimes, delivery of **punishments** for non-contribution. Just think of the tax collection system.

- To study whether punishments can enforce cooperation, consider the VCM with one modification. There is a second stage of the game in which players can mete out punishments by reducing other players' payoffs at a marginal cost of $c \in (0, 1)$. In particular, the ultimate payoff of player i is given by

$$x_i(g_1, ..., g_n, p_1, ..., p_n) = y - g_i + a \sum_{j=1}^{n} g_j - \sum_{j=1}^{n} p_j^i - c \sum_{j=1}^{n} p_i^j,$$

where p_i^j is the punishment imposed by player i on player j. Under self-regarding preferences, nobody will punish in the second stage ($p_i^j = 0$ for all i and j), because punishments are privately costly. Hence the unique subgame-perfect equilibrium is for all the players not to punish at all and to contribute nothing ($g_i = 0$ for all i).

Fehr and Gachter (2000)

- The authors implement VCM with punishments in the lab. The main objective is to investigate the impact of the punishment institution on contributions and free-riding in VCM.

- Experimental design:

TABLE 1—TREATMENT CONDITIONS

	Stranger-treatment Random group composition in each period (Sessions 1–3)	Partner-treatment Group composition constant across periods (Sessions 4 and 5)
Without punishment (ten periods)	18 groups of size n	10 groups of size n
With punishment (ten periods)	18 groups of size n	10 groups of size n

- In each stranger treatment, there are 24 participants randomly rematched into groups of 4 in each of 10 rounds. In each partner treatment, 20 participants are split into six groups of 4 participants that stay fixed for the entire duration of 10 rounds. The design is within-subject in that all subjects participate both in the no-punishment and in the punishment condition. Ordering of treatments is balanced across sessions to control for order effects. Therefore in all treatments $n = 4$. Also, in all treatments $a = MPCR = 0.4$ and $y = 20$.

- The particular payoff structure for player i is given by

$$\pi_i = \pi_i^1 \left[1 - 0.1 \min\left\{\sum_{j \neq i} p_j^i, 10\right\}\right] - \sum_{j \neq i} c(p_i^j),$$

where π_i^1 is the first-period payoff of player i given by

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

That is, each unit of punishment reduces the first-period payoff of the punished by 10%, up to the floor of 0. The cost of punishment is given by

TABLE 2—PUNISHMENT LEVELS AND ASSOCIATED COSTS FOR THE PUNISHING SUBJECT

Punishment points p_i^j	0	1	2	3	4	5	6	7	8	9	10
Costs of punishment $c(p_i^j)$	0	1	2	4	6	9	12	16	20	25	30

- Results for the stranger treatment:

TABLE 3—MEAN CONTRIBUTIONS IN THE STRANGER-TREATMENT

	Mean contribution in all periods		Mean contribution in the final periods	
Sessions	Without punishment opportunity	With punishment opportunity	Without punishment opportunity	With punishment opportunity
1	2.7	10.9	1.3	9.8
	(5.2)	(6.1)	(4.3)	(6.8)
2	4.0	12.9	2.3	14.3
	(5.7)	(6.4)	(4.3)	(5.0)
3	4.5	10.7	2.0	13.1
	(6.0)	(4.9)	(3.8)	(4.0)
Mean	3.7	11.5	1.9	12.3
	(5.7)	(5.9)	(4.1)	(5.6)

Notes: Numbers in parentheses are standard deviations. Participants of Sessions 1 and 2 first played the treatment with punishment opportunities and then the one without such opportunities. Participants of Session 3 played in the reverse order.

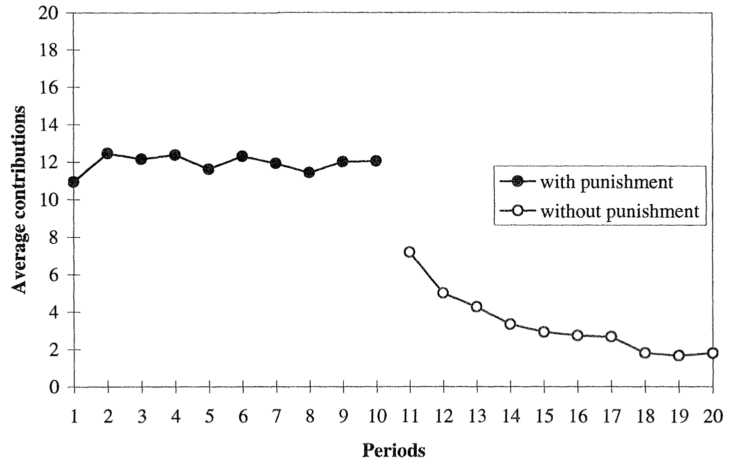

Figure 1A. Average Contributions over Time in the Stranger-Treatment (Sessions 1 and 2)

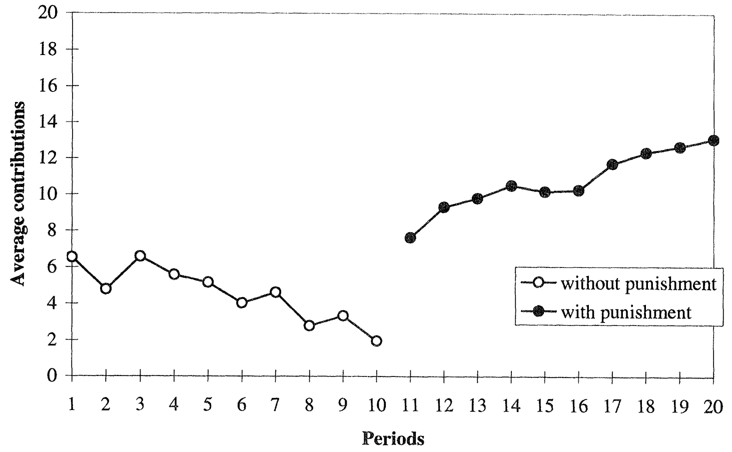

Figure 1B. Average Contributions over Time in the Stranger-Treatment (Session 3)

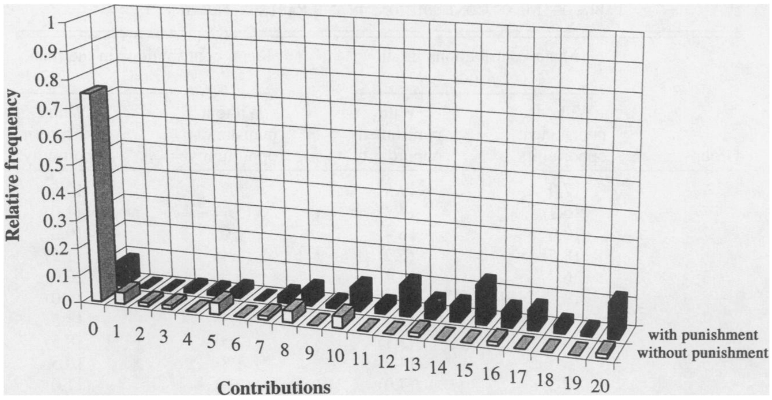

Figure 2. Distribution of Contributions in the Final Periods of the Stranger-Treatment with and without Punishment

- We observe that:

    1. The existence of punishment opportunities causes a large rise in the average contribution level in the Stranger-treatment. On average, contribution rates amount to 58 percent of the endowment.
    2. In the no-punishment condition of the Stranger-treatment, average contributions converge close to full free-riding over time. In contrast, in the punishment condition average contributions do not decrease or even increase over time.
    3. In the Stranger-treatment with punishment no stable behavioral regularity regarding individual contributions emerges, whereas in the no-punishment condition full free-riding emerges as the focal individual action.

- Results for the partner treatment:

TABLE 4—MEAN CONTRIBUTIONS IN THE PARTNER-TREATMENTS

Groups	Mean contributions in all periods: Without punishment opportunity	Mean contributions in all periods: With punishment opportunity	Mean contributions in the final periods: Without punishment opportunity	Mean contributions in the final periods: With punishment opportunity
1	7.0	17.5	5.8	19.5
	(6.3)	(4.3)	(5.1)	(1.0)
2	10.6	16.4	1.0	19.3
	(8.5)	(5.2)	(1.4)	(1.5)
3	6.7	18.4	6.3	20.0
	(7.8)	(3.6)	(9.5)	(0.0)
4	5.1	12.1	1.3	13.5
	(6.3)	(7.1)	(2.5)	(8.5)
5	6.4	14.3	1.8	10.5
	(7.2)	(7.0)	(2.9)	(11.0)
6	7.9	19.0	3.5	20.0
	(5.7)	(2.8)	(5.7)	(0.0)
7	7.4	19.0	2.5	20.0
	(7.1)	(3.4)	(2.9)	(0.0)
8	10.0	17.2	5.0	20.0
	(6.6)	(4.3)	(6.0)	(0.0)
9	3.9	17.0	0.0	20.0
	(5.9)	(5.0)	(0.0)	(0.0)
10	10.0	19.0	5.0	19.5
	(6.6)	(2.1)	(8.0)	(1.0)
Mean	7.5	17.0	3.2	18.2
	(6.8)	(4.5)	(4.4)	(2.3)

Notes: Numbers in parentheses are standard deviations. Groups 1–4 (Session 4) first played the punishment condition and then the no-punishment condition. Groups 5–10 (Session 5) played in the reverse order.

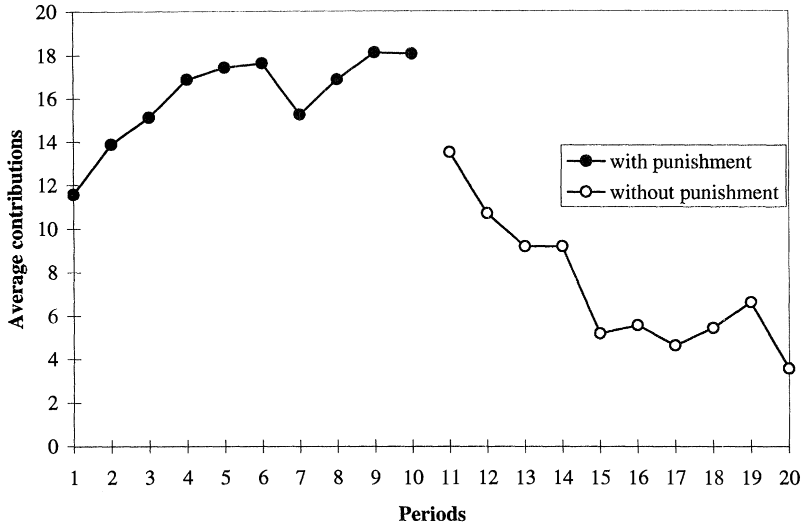

FIGURE 3A. AVERAGE CONTRIBUTIONS OVER TIME IN THE PARTNER-TREATMENT (SESSION 4)

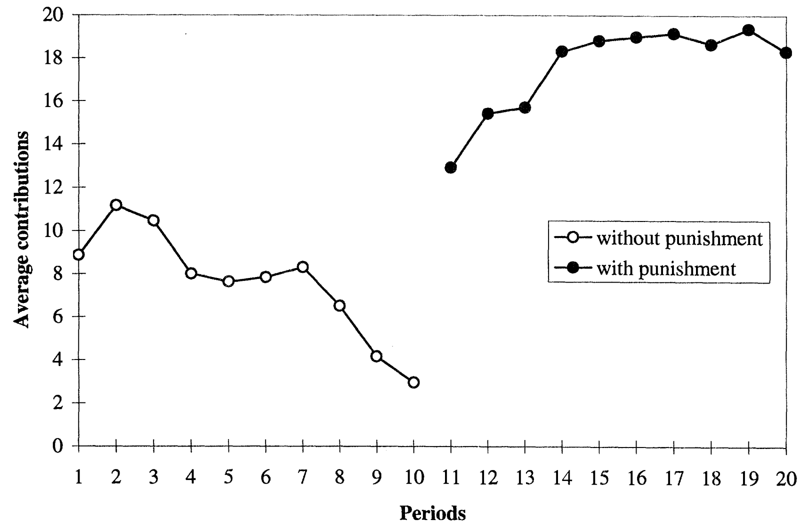

FIGURE 3B. AVERAGE CONTRIBUTIONS OVER TIME IN THE PARTNER-TREATMENT (SESSION 5)

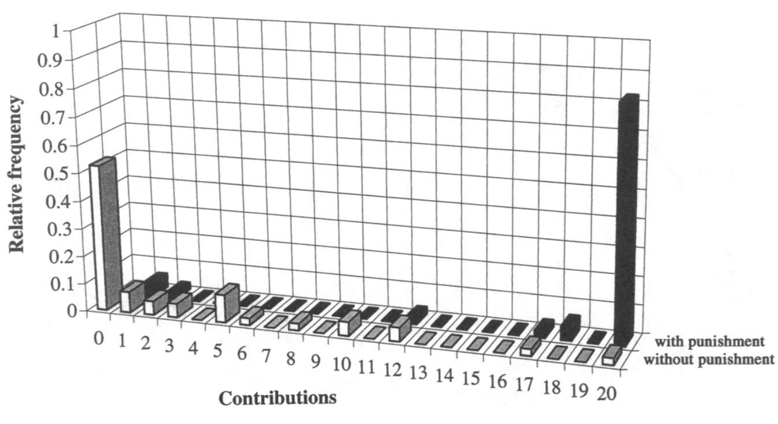

FIGURE 4. DISTRIBUTION OF CONTRIBUTIONS IN THE FINAL PERIODS OF THE PARTNER-TREATMENT WITH AND WITHOUT PUNISHMENT

- We observe that:

    1. The existence of punishment opportunities also causes a large rise in the aver- age contribution level in the Partner-treatment.
    2. In the no-punishment condition of the Partner-treatment average contributions converge toward full free-riding, whereas in the punishment condition they increase and con- verge toward full cooperation.
    3. In the Partner-treatment with punishment, full cooperation emerges as the dominant behavioral standard for individual contributions, whereas in the absence of punishment opportunities full free-riding is the focal action.

- Results on the size of the punishment:

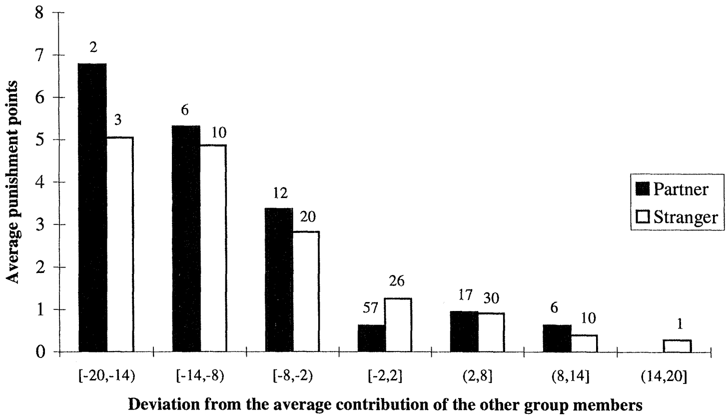

Figure 5. Received Punishment Points for Deviations from Others' Average Contribution

- We observe that:

    – In the Stranger- and the Partner- treatment, a subject is more heavily punished the more his or her contribution falls below the average contribution of other group members. Contributions above the average are punished much less and do not elicit a systematic punishment response.

- Impact on payoffs:

    – In both the Stranger- and the Partner-treatment the punishment opportunity initially causes a relative payoff loss. Yet, toward the end there is a relative payoff gain in both treatments. In particular, in the Stranger-treatment the relative payoff gain of the punishment condition is positive in the last two periods, whereas in the Partner- treatment it is positive from period 4 onward. In the final period the relative payoff gain is roughly 20 percent in the Partner-treatment and 10 percent in the Stranger-treatment.

Gureck, Irlenbusch and Rockenbach (2006)

- Once an institution such as punishment of non-cooperators is exogenously imposed on a society, contributions increase, but it is not clear how did such an institution come into existence in the first place. The authors therefore experimentally investigate how such institutions come into being.

- The authors are interested in evolutionary survival of the sanctioning institution (SI) and the sanction-free institution (SFI). They implement the linear public goods game with 7 sessions of 12 subjects and 30 rounds. Each round progresses as follows:

  1. **Stage S0**: each participant chooses whether he or she wants to belong to the group that uses the SI or the SFI; the former institution allows both rewards and punishments
  2. **Stage S1**: linear public goods game with in which each player has the endowment of 20 points; each contributed point benefits the group account with 1.6 points; hence the MPCR is $1.6/n$, where n is the number of subjects joining the given institution
  3. **Stage S2**: each member of the SI can assign between 0 and 20 points in total to other members of his/her institutional group; a punishment point hurts the punished individual 3 points, whereas a reward point benefits the rewarded individual 1 point
  4. **Feedback**: everyone receives a detailed anonymous feedback about all other players, their actions and payoffs

- Results and story:
  - In period 2, there is an imitation of free-riders in SFI since they have the highest payoffs. Over time, this leads to the reduction of contributions and payoffs in SFI, as observed in previous experiments.
  - Comparing payoffs of two dominant behavioral patterns, high contributors in SI (contributing 15 or more) and free-riders in SFI (contributing 5 or less), reveals that the former do better since period 5 onwards.

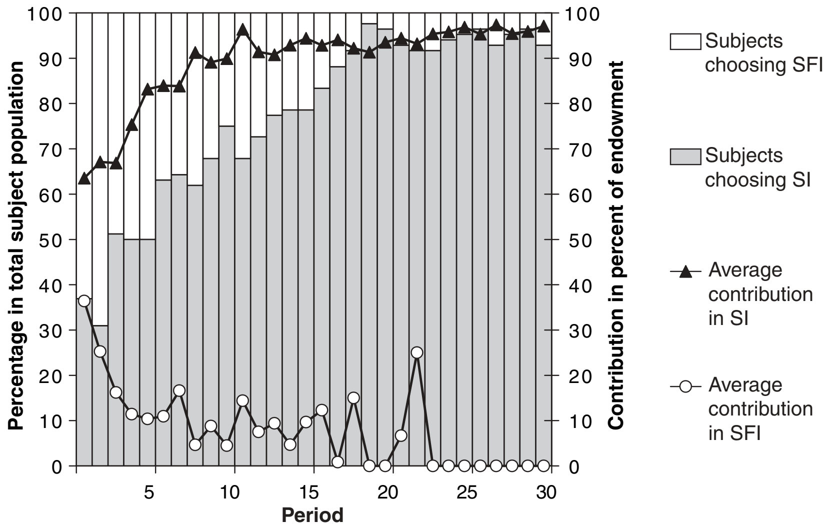

Fig. 1. Subjects' choice of institution and their contributions. The average contributions in both institutions over the 30 periods of the interaction are measured as the percentage of endowment contributed to the public good.

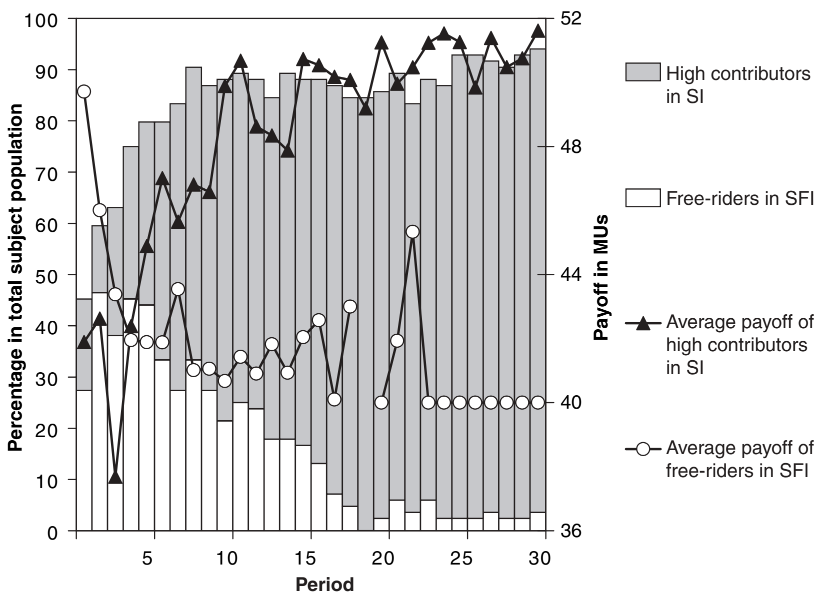

Fig. 2. Payoffs of the two predominant behavioral patterns, "free-riders" (contributions between 0 and 5 MUs) in the sanction-free institution (SFI) and "high contributors" (contributions between 15 and 20 MUs) in the sanctioning institution (SI). The highest attainable payoff (under full contributions of all subjects and no punishment) is 52 MUs and the payoff from complete free-riding and no punishment is 40 MUs.

- This causes that payoff-maximizing agents to switch from free-riding in SFI to contributing in SI over time.
- This has an amplifying effect in SI in that the payoff of contributors grows with the number of contributors in the group, attracting further converts.

Figure S1

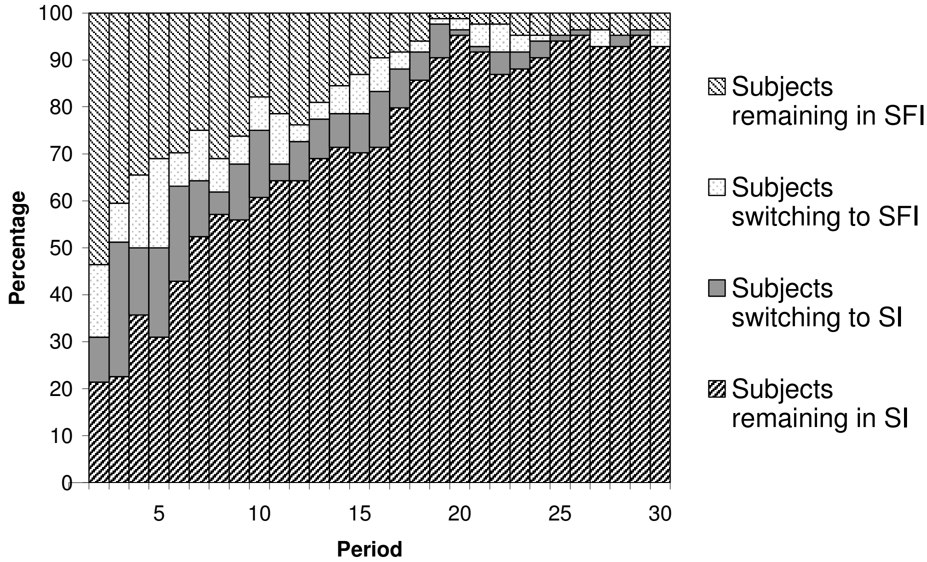

Fig. S1. Subjects' choices of institutions and their switching behavior in both directions.

Figure S2

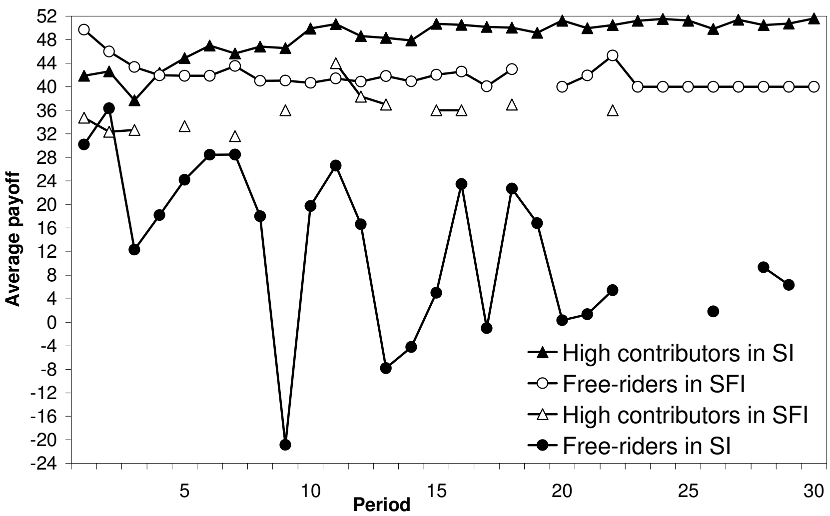

Fig. S2. Subjects' average payoffs dependent on their contribution behavior.

– At the individual level, switchers from SFI to SI typically dramatically increase their contributions, whereas switchers in the opposite direction (very rare) dramatically decrease their contributions.

– Although providing punishment is a second-order public good, about two thirds of the switchers to SI start punishing immediately.

– There is a stable proportion of high contributors and punishers (40 to 50 percent) in the SI over time.

– In SI, the payoff difference between high contributors and non-punishers and high contributors and punishers diminishes over time as the execution of punishment becomes more and more rare and as the punishment duty is shared by a higher and higher number of individuals.

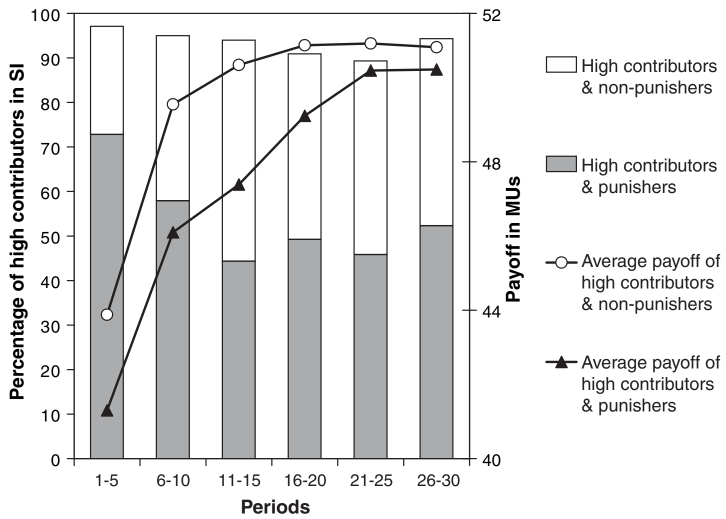

Fig. 3. Payoffs and percentages of punishers and nonpunishers among the "high contributors" (contributions between 15 and 20 MUs) in the sanctioning institution (SI). The highest attainable payoff (under full contributions of all subjects and no punishment) is 52 MUs and the payoff from complete free-riding and no punishment is 40 MUs.

– In the end, close to the most efficient allocation is achieved in SI, whereas close to the most inefficient allocation is achieved in SFI.

Kosfeld, Okada and Riedl (2009)

- The authors observe that the previous work on institution formation (Gureck et. al, 2006, as well as other papers) do not allow for the possibility of some agents free-riding on the punishment institution. This is an important limitation that the current paper tries to overcome.

- The authors are interested in evolutionary survival of a centralized sanctioning institution, called "organization," that is costly, supported by the members, and only capable of punishing the members. Others may stay out of the organization, but everyone is still in the same group. Hence non-participants may free-ride on organization members.

- The authors implement a linear public goods game with 4 players and partner matching. Each round progresses as follows:

  1. **Stage 1**: each subject chooses whether he or she wants to participate in the organization or not
  2. **Stage 2**: beliefs are elicited from each subject using quadratic scoring on how many other subjects in the group he/she expects to be interested in participation
  3. **Stage 3**: all subjects learn how many other subjects in the group want to form an organization; then each subject who expressed his/her willingness to form an organization in Stage 1 decides whether he/she indeed wants to form the organization; the organization is formed if and only if all such participants indeed decide to form the organization
  4. **Stage 4**: linear public goods game in which each player has the endowment of 20 points; each contributed point benefits the group account with 1.6 points, or, in a different treatment, 2.6 points; hence the MPCR is 0.4 (treatment IF40) or 0.65 (treatment IF65); the sanctioning mechanism confiscates any non-contributed points of the organization participants; the cost of the organization is 2 and it is equally shared by all the participants in the organization

- The authors also implemented two VCM control treatments with the same MPCRs.

- 164 subjects in total, 44 in each treatment, 40 and 36 in the two controls.

- A theoretical analysis reveals that there are two equilibria: a **status quo equilibrium** (without SI and with free-riding by everyone) and an **organizational equilibrium** (with SI formed by a threshold number of agents, contributions by these agents, and free-riding by all the non-participants).

- The idea of a threshold is parallel to the idea of the provision point. Each participant in the SI is pivotal and it is hence in his/her interest to keep supporting the institution and to fully contribute, despite the presence of free-riders.

- By introducing other-regarding preferences a'la Fehr and Schmidt (1999), one could support equilibria with universal participation in the SI, called **grand organizations**. Data on player's beliefs suggest that these are not due to miscoordination.

- Results:
  - There is almost always at least one subject in each group who wants to establish an organization. In IF40, an organization is implemented 43% of the time, in IF65 it is 61% of the time. In IF40, 83% of these organizations are grand organizations, in IF65 it is 68%. There are very few below-the-critical-threshold organizations.

TABLE 1—INITIATED AND IMPLEMENTED ORGANIZATIONS

	Treatment			
	IF40		IF65	
	Number	Percentage	Number	Percentage
Initiated organizations	220	100	216	98
Implemented organizations				
Total	95	43	132	61
One member	0	0	5	4
Two members	1	1	15	11
Three members	15	16	22	17
Four members	79	83	90	68

Notes: The table presents the absolute and relative number of *initiated* and *implemented organizations* over all rounds. Relative numbers are calculated as follows: initiated organizations relative to all rounds, implemented organizations relative to all initiated organizations, different size of organizations relative to all implemented organizations.

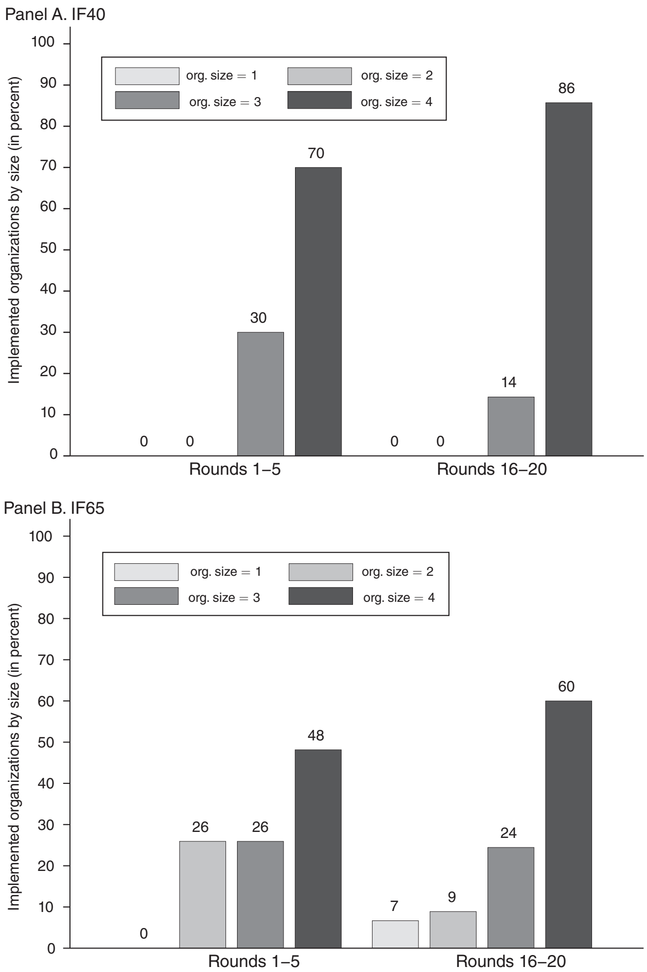

FIGURE 1. DISTRIBUTION OF IMPLEMENTED ORGANIZATIONS IN EARLY AND LATE ROUNDS

- Most of the subjects who display interest in participation in Stage 1 expect all other subjects to be interested in participation as well.
- Candidate organizations with fewer than four participants in Stage 3 have a high likelihood of being rejected. Only grand organizations have a substantial likelihood of being implemented. Overall, the likelihood of implementation increases with the MPCR.

TABLE 2—BELIEFS AND RATE OF IMPLEMENTATION

	Treatment									
	IF40					IF65				
		Number of participants					Number of participants			
Belief	Observations	1	2	3	4	Observations	1	2	3	4
First round	26	9.42	18.19	34.50	37.88	25	19.52	14.48	23.80	42.20
Final round	35	5.29	4.83	15.86	74.03	32	2.34	5.47	21.28	70.91
All rounds	726	6.48	7.03	21.67	64.81	671	5.01	11.80	21.75	61.44
		Number of participants					Number of participants			
Implementation rate		1	2	3	4		1	2	3	4
All rounds		0.00	2.94	23.08	69.30		27.78	37.50	37.29	90.91
Observations		7	34	65	114		18	40	59	99

Notes: The upper panel of the table presents the average probability belief (in percent) of participating players in stage one of the game about the total number of participants in the organization. The lower panel presents the likelihood of implementation (in percent) of an organization depending on the number of participating players.

- Overall, the possibility to form the organization has a positive impact on contributions to the public good, their stability and overall efficiency.

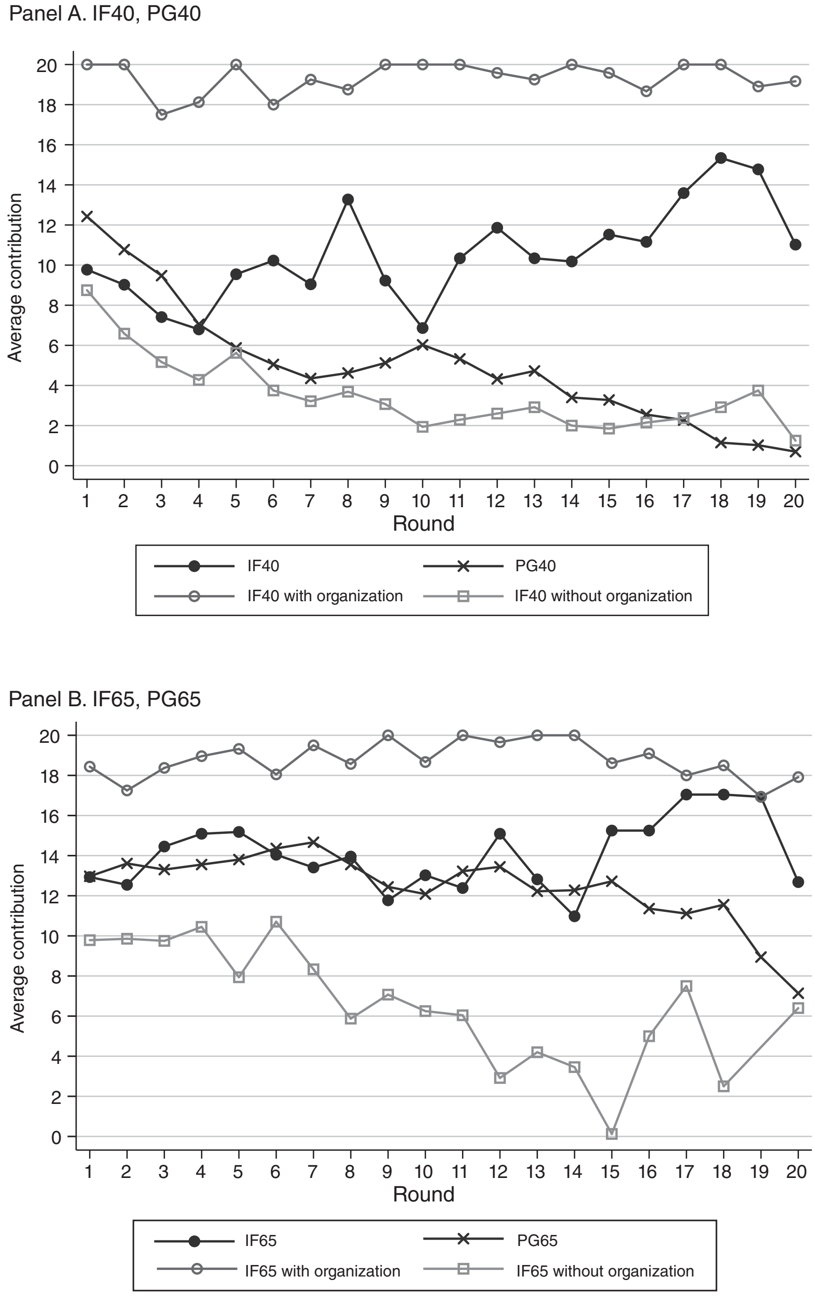

Figure 2. Average Contribution to the Public Good with and without the Possibility of Institution Formation

A GUIDE TO FURTHER READING

- **Isaac and Walker (1988b)** show that, in VCM, costless communication about contributions between rounds increases contributions, and this increase sometimes persists even after the possibility to communicate is removed.

- **Andreoni (1993)** presents an experimental test of the proposition that government contributions to public goods, financed by lump-sum taxation, will completely crowd out voluntary contributions. He finds that crowding-out is incomplete and that subjects who are taxed are significantly more cooperative. This suggests that people experience some benefit (warm glow) from contributing to public goods.

- **Andreoni (1995)** investigates why subjects cooperate in VCM. Particularly, he wants to distinguish between two leading hypotheses: (1) kindness/altruism/warm-glow; (2) errors/confusion. He finds that about half of all cooperation comes from subjects who understand free-riding but choose to cooperate out of some form of kindness, whereas the other half is due to errors/confusion.

- **Bagnoli, Ben-David and McKee (1992)** consider the provision point mechanism when multiple units of the public good can be provided and players have downward-sloping demands. They find that public good is provided at an efficient level only in about half of the cases, the welfare level reaches only about two thirds of its potential. These results are much weaker then the corresponding results for the discrete provision case analyzed by Bagnoli and McKee (1991).

- **Rondeau, Poe and Schulze (2005)** compare the efficiency of VCM and provision-point mechanism using meta-analysis. They find that, overall, the PPM increases total contributions, it is more responsive to changes in induced value, and is generally more efficient than the VCM. For public goods with a benefit–cost ratio in the interval [1, 1.4), however, the VCM captures a greater portion of available benefits than the PPM.