Cooperation by Evolutionary Feedback Selection in Public Good Experiments

(Genetic Evolution, norm internalization and Nash learning)

Didier Sornette (Chair of Entrepreneurial Risks)

Collaboration with Didier Darcet (Renaissance Investment Management)
Moritz Hetzer (PhD student at ETHZ) and Paolo. Vanini (ZKB)

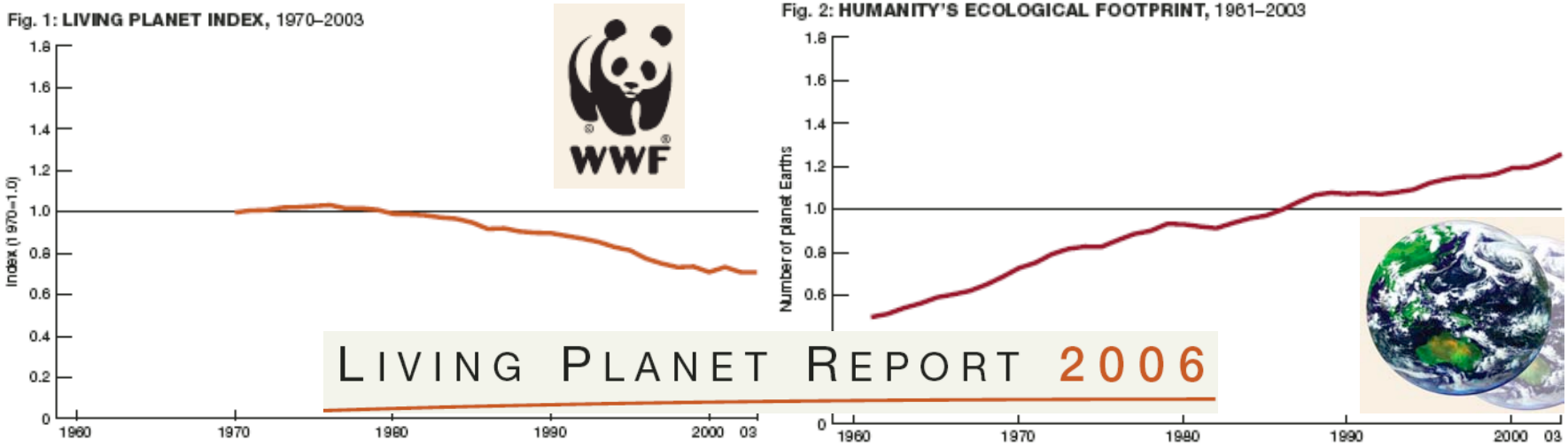
MOTIVATIONS

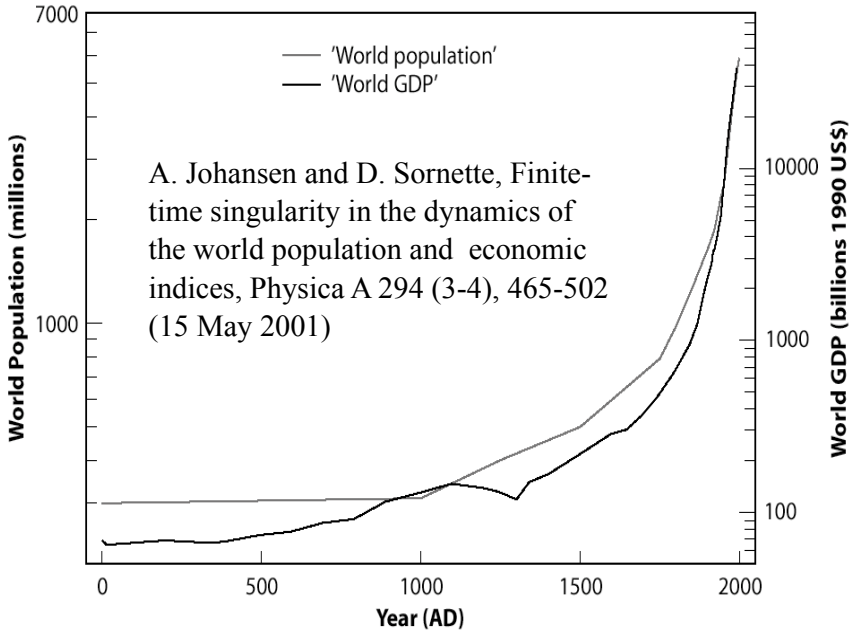
- why do we cooperate? Why don't we cooperate enough? Social order, collective action, feedback of society/culture on individuals' preference

- Economics and the Social Sciences to include "selfish" + "other-regarding" behaviors

- Sustainability of Mankind

Combining environmental and social sciences for decisions



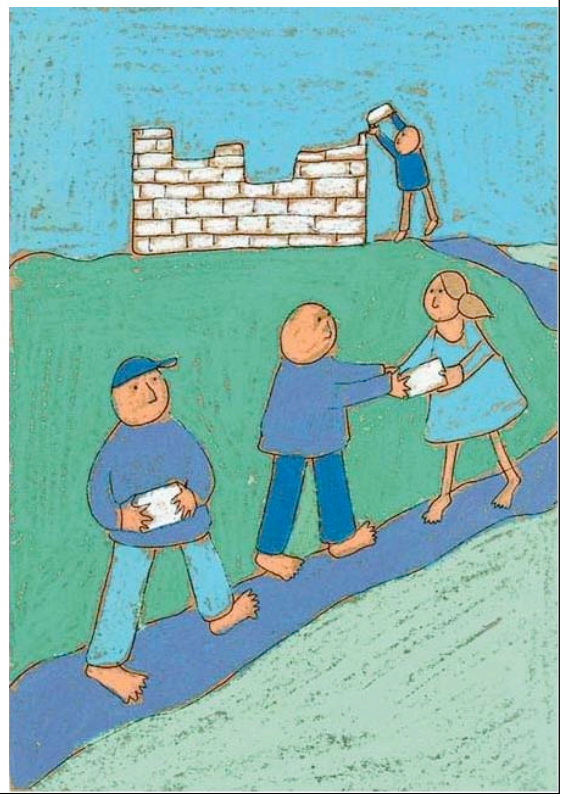


A. Johansen and D. Sornette, Finite-time singularity in the dynamics of the world population and economic indices, Physica A 294 (3-4), 465-502 (15 May 2001)

COOPERATION

The potential for cooperation is everywhere in nature, yet evolution seems to rarely take advantage of it.

When it does - social insects, multi-cellular organisms, human societies - the results can be spectacularly successful.



One of the most striking characteristics of Homo sapiens is our sociality and cooperation.

“Standard” theories explaining animal cooperation

(1) Evolutionary kin selection (Hamilton, 1964) roots behavior to genetic relatedness using Hamilton’s notion of inclusive fitness.

(2) Costly signaling theory (Zahavi, 1977; Gintis et al., 2001) and indirect reciprocity (Nowak and Sigmund, 1995; 1998; Leimar and Hammerstein, 2001) base cooperation in large groups on the build-up of the reputation of cooperators.

(3) Reciprocal altruism or direct reciprocity (Trivers, 1971; Axelrod, 1984; Nowak et al., 1995)

-derives cooperation from selfish motives in the presence of long-term repeated interactions.

-partnership or delayed reciprocity: help now to prevent the loss of a future helper (Eshel and Shaked, 2000)

HUMAN COOPERATION

A growing number of experiments (Fehr and Fischbacher, 2003) have documented behaviors which seem incompatible with the self-regarding nature of humans involved in these mechanisms (1-4):

Humans show a predisposition to cooperate in unrepeated interactions with strangers at their own cost in absence of witnesses.

(4) Strong Reciprocity (Gintis et al., 2005)
"altruistic rewarding" (Fehr and Fischbacher, 2003)
"altruistic punishment" (Fehr and Gächter, 2002)

Potential for group selection and culture evolution (Richerson and Boyd, 2005) above the individual level as a mechanism for spontaneous cooperation within human groups, because of the competitive advantage such cooperation affords.

HUMAN COOPERATION

But, at the same time, the group face the problems of

-internal competition between individuals (or subgroups) and

-the free-rider, to which altruistic punishments/rewards constitute a possible remedy (but second-order problem…).

The theory of strong reciprocity posits that humans are still self-centered but also inequity adverse, that is, their utility function includes terms sensitive to unfairness.

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

Subjects suffer more from inequity that is to their material disadvantage than from inequity that is to their material advantage. $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$.

Fehr and Schmidt (1999)

“Social preference”

QUESTIONS

- What is the origin of “inequity aversion”?
- What is its relative strength vs self-centered interest?
- How does it reveal itself in different contexts?

EVOLUTIONARY FEEDBACK SELECTION

•It is well-known that altruistic punishments and rewards can emerge from evolution.

•Our proposal: evolution has chiseled our abilities to adapt our altruistic punishments and rewards in human beings to an "optimal" level dependent on context.

•The level of cooperation and punishment is AS IF an optimal cost-benefit analysis is performed on the part of the each agent.

•The individual preferences are resulting from a collective global organization at the level of the group, rather than an idiosyncratic optimization.

EVOLUTIONARY FEEDBACK SELECTION

•The theory of evolutionary feedback selection allows us to explain quantitatively the experimental results on

-third-party punishment games,
-the ultimatum game
-altruistic punishment games.

•Numerical simulations of a simple evolutionary agent-based model of repeated agent interactions with feedbacks:

-supports evolutionary feedback selection,
-suggests that the propensity to punish can be seen as an emergent property (Anderson, 1972; Goldenfeld and Kadanoff, 1999).

•The theory of evolutionary feedback selection emphasizes an evolutionary mechanism, in which the trait of providing certain levels of feedbacks is selected for and co-evolves with an increasing individual fitness, which is mediated by the collective output.

General formulation of the theory of evolutionary feedback selection

Let us consider n agents in a social dilemma situation, in which

-voluntary contributions are needed to obtain some shared end-result,

-the individual rational choice is to free-ride.

We denote by c_i i=1, …, n, the contributions of each agent to the common project.

Then, the shared end-result is quantified by n payoff functions $P_i(c_1, c_2, \ldots, c_i, \ldots, c_n)$, which are possibly distinct.

General formulation of the theory of evolutionary feedback selection

After the one-period play, the total wealth of agent i is therefore

$$P_i(c_1, c_2, \ldots, c_i, \ldots, c_n) - c_i.$$

The conditions for a social dilemma to hold is that

(1a) $P_i(0, 0, \ldots, 0, c_i, 0, \ldots, 0) < c_i$,
(no individual incentive)

(1b) $P_i(c_1, c_2, \ldots, c_{i-1}, 0, c_{i+1}, \ldots, c_n) > P_i(c_1, c_2, \ldots, c_i, \ldots, c_n) - c_i$
(free ride)

(1c) $P_i(c, c, \ldots, c, c, c, \ldots, c) - c > 0$
(collective cooperation rewards the individual)

Feedback by reward and punishment

(i) Punishment of the free rider:

<u>the contributions are different but the payoffs are identical</u>, a situation often encountered in global resource sharing, such as for instance in public good interactions and in the tragedy of the commons.

The level of punishment exerted by an agent k on an agent i observing the contributions c_i and c_j of two agents i and j (k can also be j herself) should be an increasing function of $|c_i - c_j|$, with the punishment applied to the smallest contributor.

$$Pu_{k\rightarrow i} = \sum_{[j=1 \text{ to } n]} \kappa_{kji}^{c}(c_j - c_i)$$

$\kappa_{kji}^{c}(c_j - c_i)$ and $\kappa_{kji}^{P}(P_i - P_j)$ are non-decreasing functions

Feedback by reward and punishment

(ii) Punishment of the injustified/unfair endowment:
<u>the contributions are identical but the payoffs are different</u> (random factors or some structural asymmetry). This situation is characteristic of the ultimatum and dictator games for instance.

The level of punishment exerted by an agent k (on agent i) observing the payoffs P_i and P_j of the two agents i and j should be an increasing function of $|P_i - P_j|$ with the punishment applied to the greatest payoff.

$$Pu_{k \to i} = \boxed{} \sum_{[j=1 \text{ to } n]} \kappa_{kji}^{P}(P_i - P_j).$$

$\kappa_{kji}^{c}(c_j - c_i)$ and $\kappa_{kji}^{P}(P_i - P_j)$ are non-decreasing functions

(iii) A mixture of (i) and (ii) in which both contributions and payoffs are different.

$$\boxed{Pu_{k\rightarrow i} = \sum_{[j=1 \text{ to } n]} \kappa_{kji}^{c}(c_j - c_i) + \sum_{[j=1 \text{ to } n]} \kappa_{kji}^{P}(P_i - P_j).}$$

Punishment of free rider

Punishment of unfair contribution

Hypothesis: "Self" or available "consensus" act as the reference point

$\kappa_{kji}^{c}(c_j - c_i) = \kappa_{kji}^{c} * (c_j - c_i)$ for $c_j - c_i > 0$ and 0 otherwise.

$\kappa_{kji}^{P}(P_i - P_j) = \kappa_{kji}^{P} * (P_i - P_j)$ for $P_i - P_j > 0$ and 0 otherwise.

Simplification: $Pu_{k\rightarrow i} = \kappa^{c}(n-1)(\langle c_j \rangle_i - c_i) + \kappa^{P}(n-1)(P_i - \langle P_j \rangle_i)$,

Hypothesis:

People

(i) are aware that they can profit from cooperation,

(ii) can profit even more if they are among a minority of non-cooperators with a majority of cooperators,

(iii) know that they may be punished if they behave unfairly,

(iv) possess the drive to punish other agents who are unfair from their perspective.

These elements are part of the inputs and brain processing abilities that human beings developed over many generations of cooperation/defection experiences.

The level of cooperation of a given agent i is assumed to be determined by the following optimization problem that a given agent i must solve to determine her contribution c_i

$$\text{Max}_{c_i} \; E_i[\, P_i(c_1, c_2, \ldots, c_i, \ldots, c_N) - c_i - \sum_{[k=1 \text{ to } n]} Pu_{i \to k} - r_p \sum_{[j=1 \text{ to } n]} Pu_{j \to i} \,]$$

r_p is the so-called reward/punishment efficiency parameter

The optimization problem for agent i consists in determining her optimal contribution c_i based on her expectation of the contributions of the other agents, which determines her expectation of the punishment she may have to endure from other agents and the punishment she may be inclined to impose on the other agents.

What are the specificities of our approach?

(i) Seeing fairness as resulting from the propensity to punish in the presence of unfairness (rather than being an innate contribution to one's utility),
(ii) Punishment is NOT necessarily to re-establish fairness,
(iii) Quantitative determination of the level of punishment.

The main target of our theory is to determine endogenously the quantitative value of the propensity to punish and the level of cooperation by a condition of maximum selfish gain in an evolutionary stable collective equilibrium point.

Analysis of a third-party punishment game

(Fehr and Fischbacher, 2003)

-Allocator is endowed with 100 MUs and is given the free option to donate part m_a

-Recipient

-Third party, who receives 50 MUs, can spend money to punish the allocator, when informed of the money transfer from the allocator to the recipient.

Every MU spent on punishment by the third party reduces the allocator's wealth by 3 MUs (punishment efficiency r_p=3).

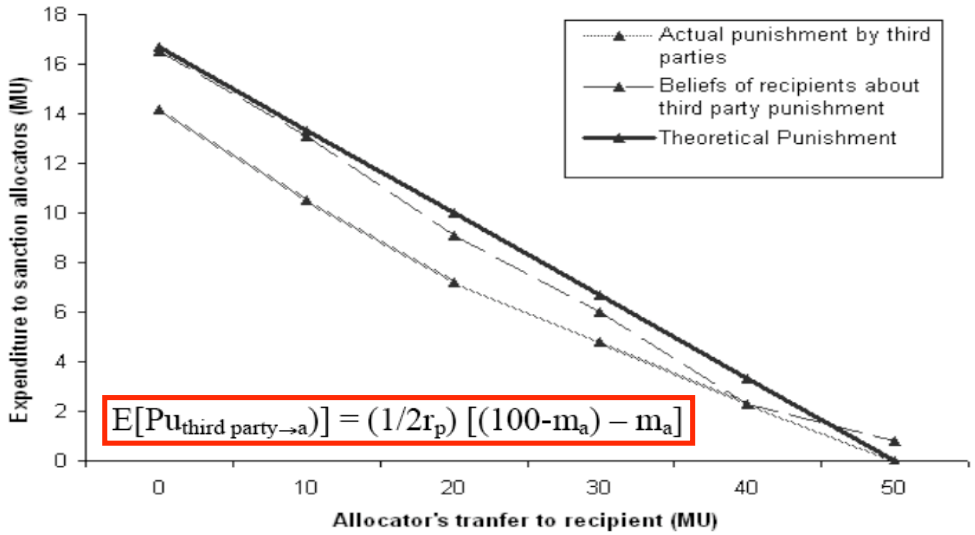


Fig.1: Mean expenditure by the punishing third party (in MUs) as a function of the transfer by the allocator to the recipient (in MUs) obtained in the experiments reported in (Fehr and Fischbacher, 2003). The theoretical punishment represented by the thick line is given by expression (9) with the value r_p=3 of the punishment efficiency used in (Fehr and Fischbacher, 2003).

The Ultimatum game

Two subjects have to agree on the division of 100 MUs.

-The proposer makes exactly one proposal $0 \leq m_a \leq 100$ to the responder and $(100-m_a)$ for him.

-The responder can accept or reject the proposed division.

-In the case of rejection, both receive nothing. In the case of acceptance, the proposal is implemented.

Probability to reject: $p_r = k\,(100-2m_a)/m_a$

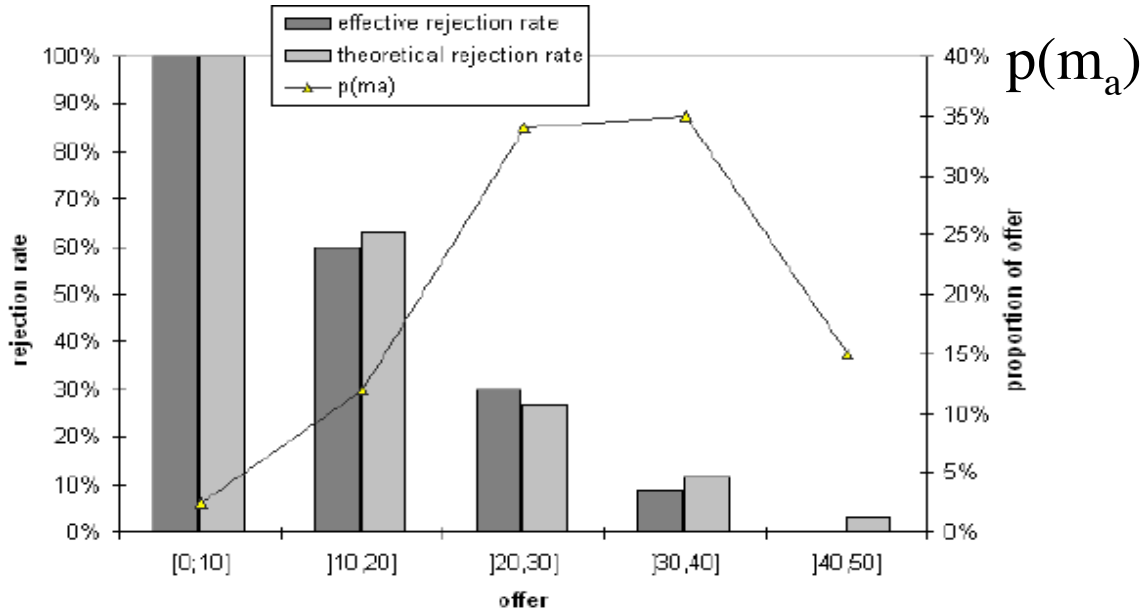
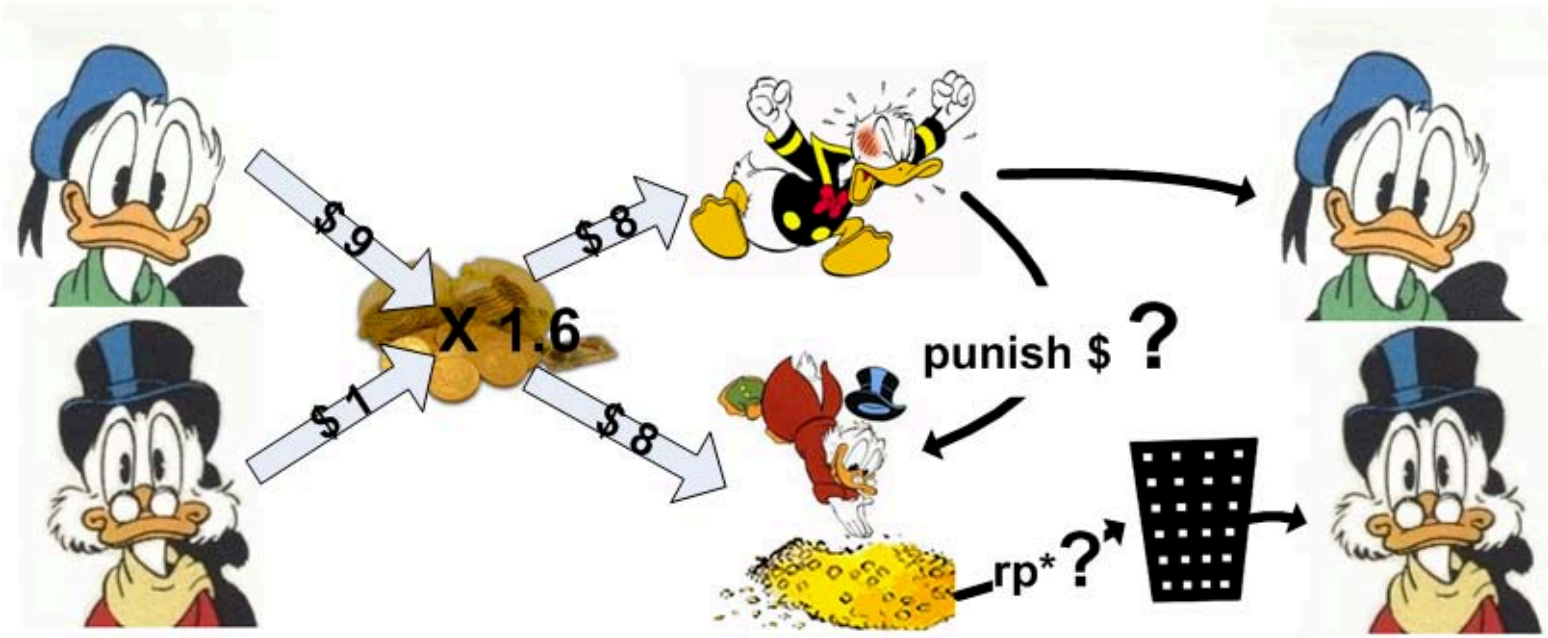


Fig.2: Experimental rejection rates in an ultimatum game involving French participants performed by Boarini, Laslier and Robin (2004) compared with theoretical rejection rates. The theoretical rejection rate is calculated using expressions (13) with (12), where the expectation in (12) is performed over the observed distribution $p(m_a)$ reported in (Boarini et al., 2004), which is shown as the continuous line joining the triangles. There is thus no adjustable parameter in the theory. Note that the cooperation reported in (Boarini et al., 2004) is stable over time.

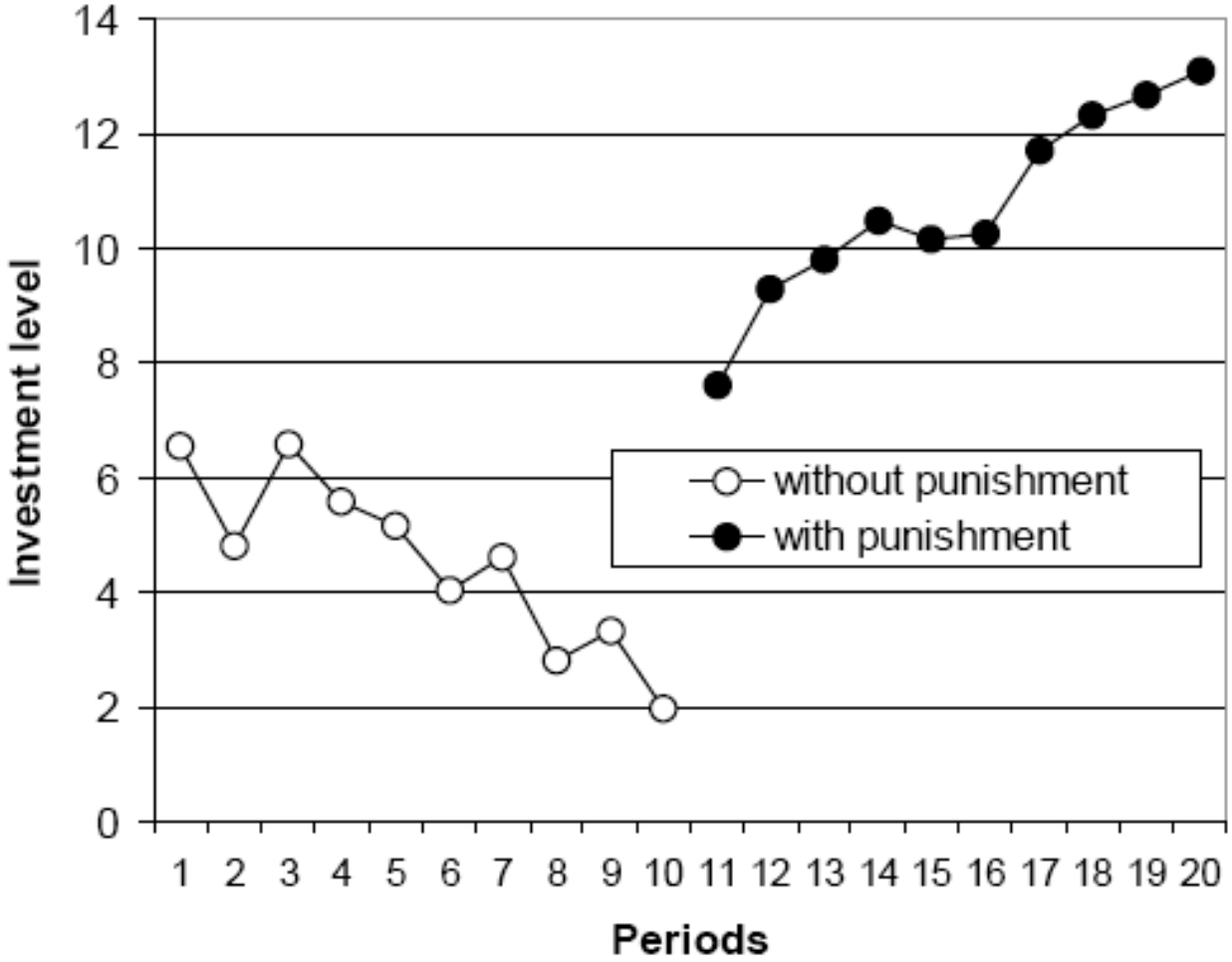
The Altruistic Punishment game

Cooperation in public good games



- **The questions are:**
  - Which amount is Donald going to spent to punish uncle Scrooge ?
  - Why does uncle Scrooge only contribute $1 and Donald $9 ?

Without feedbacks, rational decision theory predicts that cooperation should break down after a few rounds, which is indeed observed experimentally.



Fehr and Gächter (2002; 2003)

Analysis of Fehr and Gächter (2002; 2003)'s experiments on altruistic punishments (I)

$m_1, m_2, \ldots, m_{i-1}, m_{i+1}, \ldots, m_n$ are the amounts of MUs invested by members 1, 2…,i-1, i+1,…, n in the project and which are unknown to agent i.

Expected gain $E_i[G_i \mid m_i]$ for agent i in absence of punishment, and conditioned on her contribution m_i, is:

$$
\begin{aligned}
E_i[G_i \mid m_i] &= -m_i + (r_1/n)\, E[m_1+m_2+\ldots+m_n] \\
&= -m_i\,(1-r_1/n) + (r_1/n)\, E[m_1+\ldots+mi_{-1}+mi_{+1}+\ldots+m_n] \\
&= -mi\,(1-r_1/n) + (n-1)\,(r_1/n)\, E_i[m_j]\ , \quad j \neq i.
\end{aligned}
$$

(All agents are symmetric as viewed by i)

Per agent:

$$E_i[G_i \mid m_i] = \sum_{[j=1 \text{ to } n;\ j \neq i]} E_i[G_{i,j}]$$

with

$$E_i[G_{i,j} \mid m_i] = -\, a\, m_i + (r_1/n)\, E_i[m_j], \quad \text{with} \quad a = (1 - r_1/n)/(n-1)] > 0\ ,$$

Since $r_1/n < 1$, $a > 0$, incentive to defect

Agent i uses a (possibly subjective) probability $Pj_(p)$ to quantify her belief that the agent j will invest p MUs in the group project:

$$E_i[m_j] = \sum_{[p=0 \text{ à } 20]} p\, P_j(p)$$

Our starting point is, once again, the evidence that people assign and maintain a self-perspective both in action and social interactions (Vogeley et al., 2001; Vogeley and Funk, 2003) within their 'theory of mind' (Flecher et al., 1995; Gallagher and Fritch, 2003).

Simplest assumption:

$$Pu_{j\rightarrow i} = k\,(m_j - m_i), \text{ for } m_j > m_i \;; \qquad Pu_{j\rightarrow i} = 0, \text{ for } m_j \leq m_i$$

WHAT IS k?

Total expected gain of agent i in her relation with agent j:

$$E_i[G_{i,j} \mid m_i] = -\, a\, m_i + (r_1/n)\, E_i[m_j] - r_p\, E_i[Pu_{j\to i}] - E_i[Pu_{i\to j}]$$

$$E_i[Pu_{i\to j}] = k \sum_{[p=0 \text{ to } m_i]} (m_i - p)\, P_j(p)\ ; \quad E_i[Pu_{j\to i}] = k \sum_{[p=m_i \text{ to } 20]} (p - m_i)\, P_j(p)$$

Her punishment <u>to</u> others

Her punishment <u>by</u> others

Total expected marginal gain of agent i,
who hesitates between investing m_i and m_{i+1} (with i<20):

$$E_i[G_{i,j} \mid m_i+1] - E_i[G_{i,j} \mid m_i] = -a + k\,[\,r_p\,(1-F_i) - F_i]$$

where

$$F_i = \sum_{[p=0 \text{ to } m_i]} P_j(p)$$

probability perceived by agent i that another agent j will contribute less than or equally to her m_i

(1) an anticipated loss "–a" due to the equal sharing of any investment between team members,

(2) an anticipated gain $kr_p(1-F_i)$ due to the decrease of the probability of being punished by agent j

(3) an anticipated loss $-k\,F_i$ due to the increase in the probability of punishing the other agent j

Evolutionary feedback selection has selected k and $F(m_i)$

Evolutionary feedback selection has selected k and $F(m_i)$

-Hypothesis: PDF of others' contribution = PDF of expectation of oneself contribution

-Stochastic dominated limit: robust response for k is to maximize $E_i[G_{i,j} \mid m_i]$, after averaging all possible contribution levels m_i weighted by their corresponding probabilities $P_j(m_i)$.

-Since by construction the random variable F_i is uniformly distributed between 0 and 1, its average with respect to m_i weighted by $P_j(m_i)$ is always equal to ½, $\forall\ P_j(m_i)$: $E[F_i]=1/2$
(Other interpretation: median taken as best guess for the contribution of others).

$$k = 2a/(r_p-1) = 2\,(1-r_1/n)\,/\,[(n-1)\,(r_p-1)]$$

For $n=4$, $r_p=3$, and $r_1=1.6$, we obtain $k=0.2$.

Total expenditure by punishing members towards agent i:

$$Pu_{group \to i} = \sum_{[j,\ ;\ j \neq i,\ mj > mi]} Pu_{j \to I} \qquad (**)$$

$$Pu_{group \to i} = [2(1-r_1/n)/(r_p-1)](<m_{j;\ j \neq i,\ ,\ mj > mi}> - m_i) = 0.6\ (<m_{j;\ j \neq i,\ ,\ mj > mi}> - m_i),$$

for n=4, r_p=3, and r_1=1.6.

We augment our prediction (**) by providing a standard deviation.

We simulated 20'000 synthetic random games with r_p=3, and r_1=1.6 and n=4 players.

Each agent has the same punishment coefficient $k=2a/(r_p-1)=0.2$ and her contribution m_i is taken from a uniform distribution between 0 and 20.
The punishment she exerts on other agents is given by

$$Pu_{j \to i} = k\,(m_j - m_i), \quad \text{for} \quad m_j > m_i$$

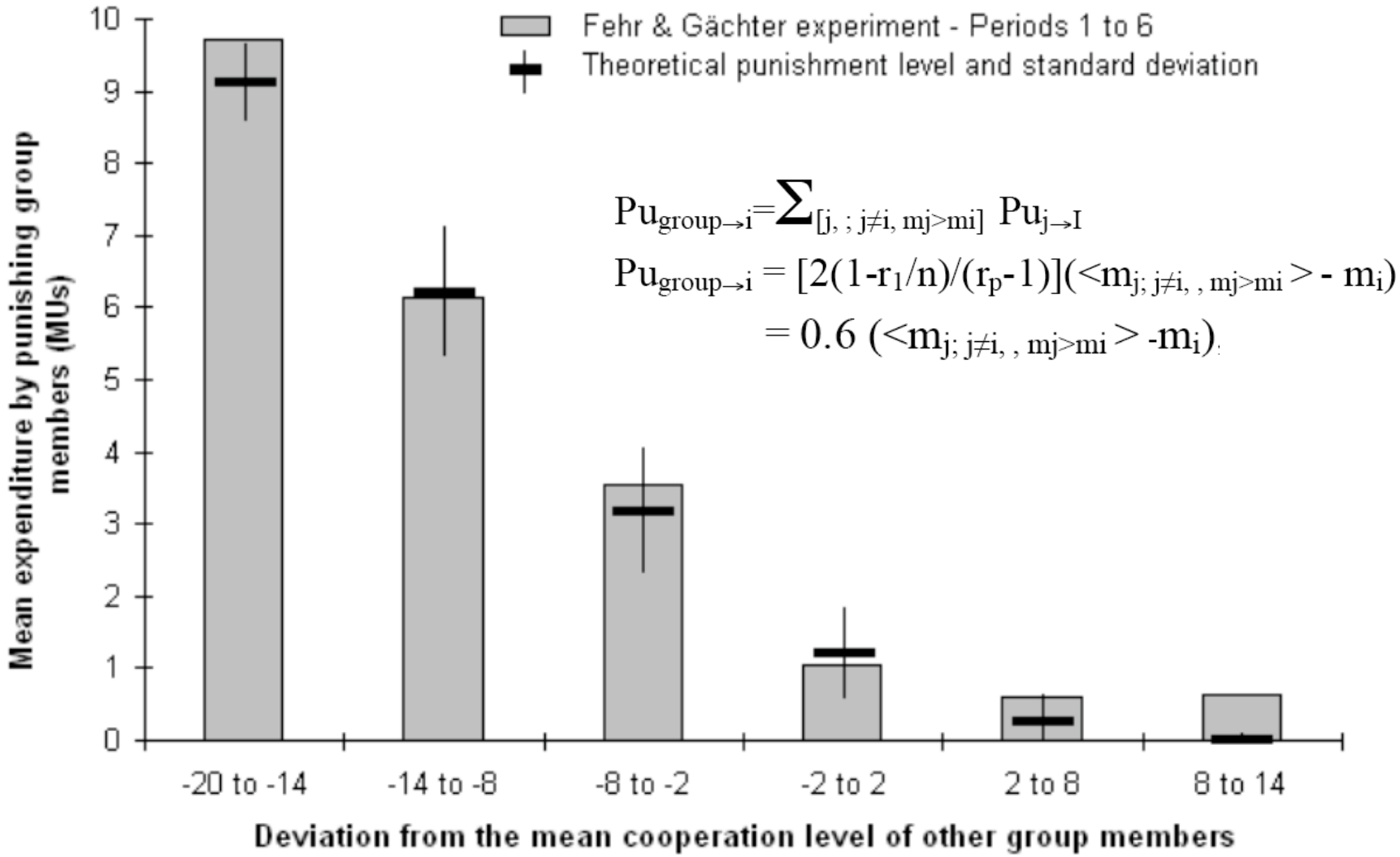


Fig.4: Mean expenditure by punishing group members (in MUs) as a function of the deviation between the contribution of an agent from the mean cooperation level of the other group members. The vertical bars are the averages of the mean expenditure over all the six periods studied by Fehr and Gächter (2002). The thick horizontal segments give the punishment level predicted by expression (**) They are accompanied by vertical segments indicating ± 1 std, obtained as described in the text.

Summary of EFS theory

$$Pu_{k\rightarrow i} = \sum_{[j=1 \text{ to } n]} \kappa_{kji}^{c}(c_j - c_i) + \sum_{[j=1 \text{ to } n]} \kappa_{kji}^{P}(P_i - P_j)$$

$$Pu_{j\rightarrow i} = k\,(m_j - m_i), \text{ for } m_j > m_i$$

Third-party game: $k \rightarrow 1/2r_p$ (payoff)

Ultimatum game: $k = E[1/(10^4/m_a^2 - 2)]$ (payoff)

Altruistic punishment game: $k = 2a/(r_p-1) = 2\,(1-r_1/n)\,/\,[(n-1)\,(r_p-1)]$

(contribution)

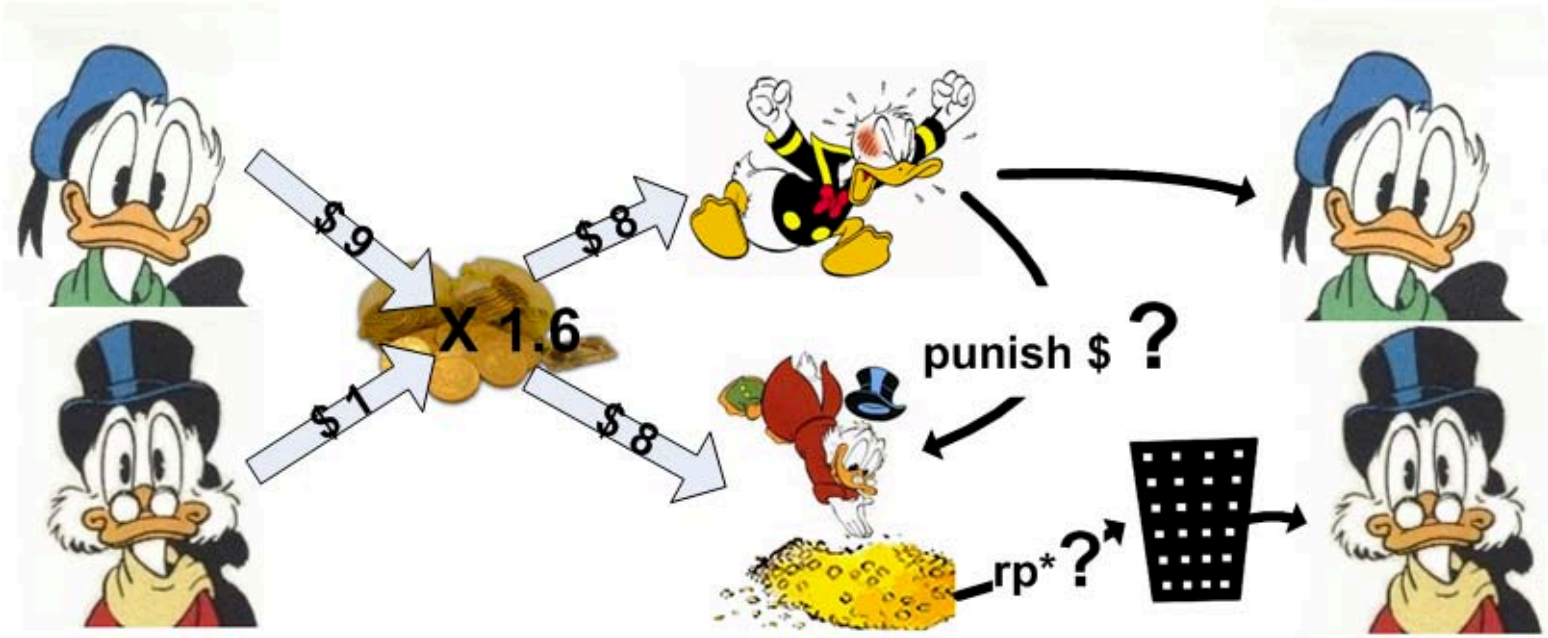
A simple ABM of repeated agent interactions with punishment

to test the plausibility of evolutionary feedback selection

(Genetic Evolution, norm internalization and Nash learning)

- The Public Goods Game
- Agent-based Model of public goods game
- ABM Results for all parameters
- Impact of group migration on the level of cooperation
- The role of consumption and wealth on the level of cooperation

Example: Cooperation in public good games



- **The questions are:**
  - Which amount is Donald going to spent to punish uncle Scrooge ?
  - Why does uncle Scrooge only contribute $1 and Donald $9 ?

Agent based model

- N agents playing the public good game.
- All are initialized with k=0, m=0, t1=1, t2=1, t3=1, wealth=50, project yield factor = 1.6 and punishment efficiency rp=3.5, where
  - **k** reflects the agents' **propensity to punish** others
  - **m** is the amount the agent **contributing** to the pool
  - **t1** reflects the threshold for the **inequality aversion** from a **self-centered gain perspective**, i.e. the agent **accepts a difference** of (1-t1)*profit/loss of her profit/loss **compared to** the **group's average profit/loss**, before she **changes** her **strategies**.

Agent based model

- Simulation steps
  1. Agents invest according to their m_i value
  2. Receive payoff from the pool
  3. Agents punish other agents according to:

$$
\begin{aligned}
Pu_{i\rightarrow j} &= k(m_i \cdot t_3 - m_j) && , m_i > t_2 \cdot m_j \\
&= 0 && , m_i \leq t_2 \cdot m_j
\end{aligned}
$$

  4. Calculating Profit/Loss

     payoff - received punishment – punishment to other – m

Agent based model

5. Evolutionary steps:

- If Profit/Loss * t1 < groups average Profit/Loss

  OR

  Profit/Loss <= 0

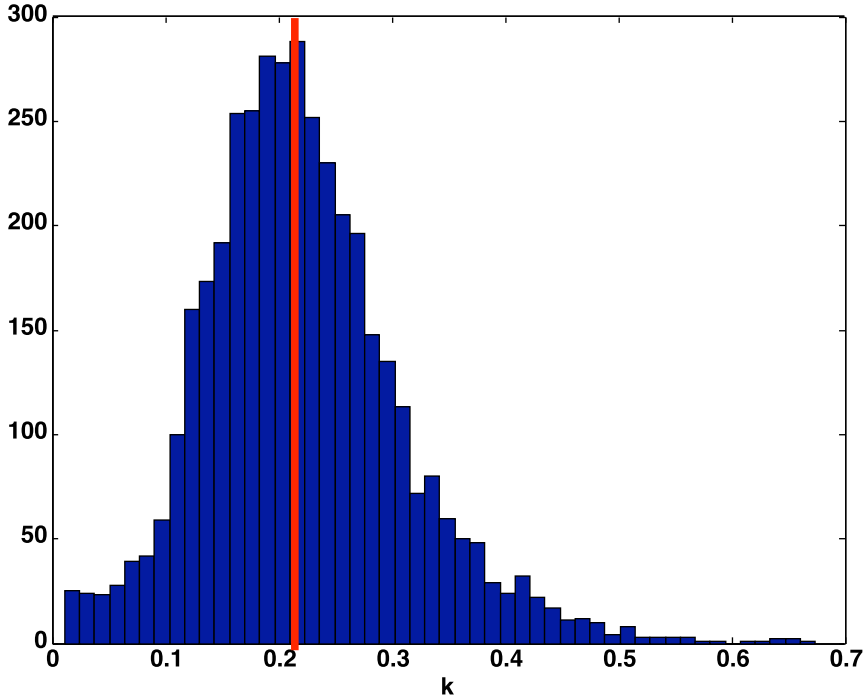
  → **Change m, k and t1,t2,t3 with a small random distributed value around their previous value**

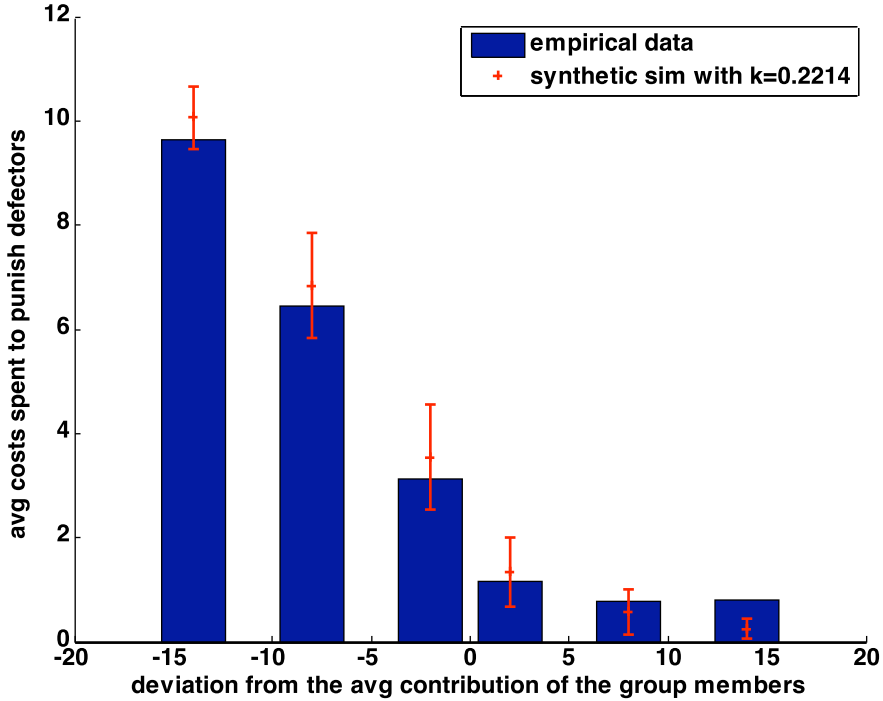
- If Wealth<0

  → **Reinitialize agent with the groups average value of k, m and t plus a i.i.d. small noise around the previous values**

ABM Results – parameter k

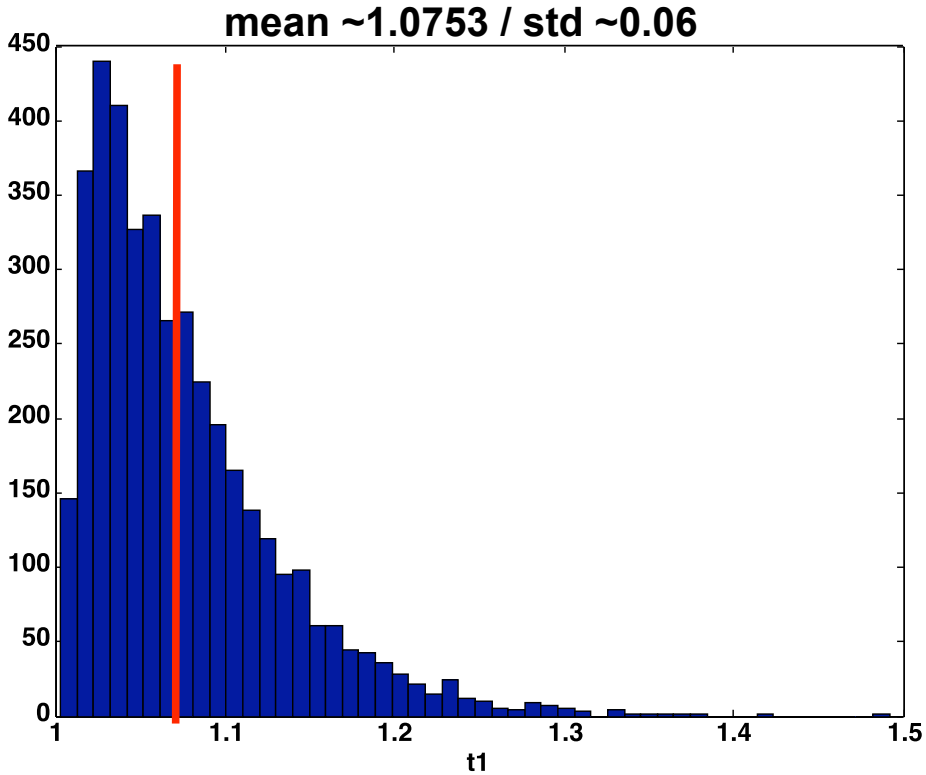
mean ~0.2214 / std ~0.09





Agents spend on average 22% of the difference between their contribution and the defectors contribution to punish the defector.

ABM Results – parameter t1



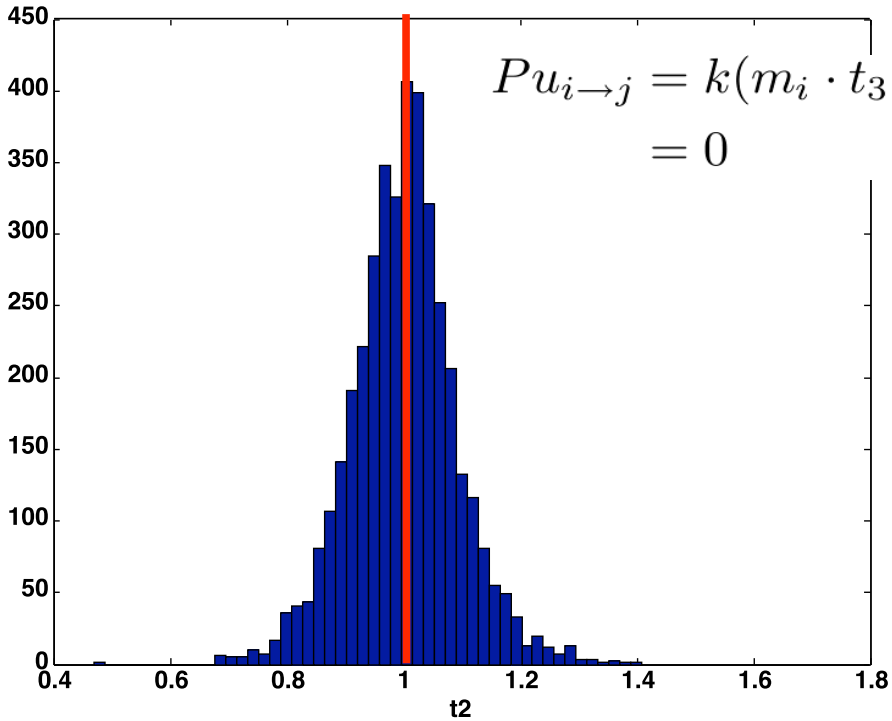
5. Evolutionary steps:

- If Profit/Loss * t1 < groups average Profit/Loss

Agents are willing to accept a difference of ~7.5% between their profit loss and the groups average profit/loss before they change their strategies (k,m,t1,t2,t3)

ABM Results – parameter t2

mean ~1.00 / std ~0.09

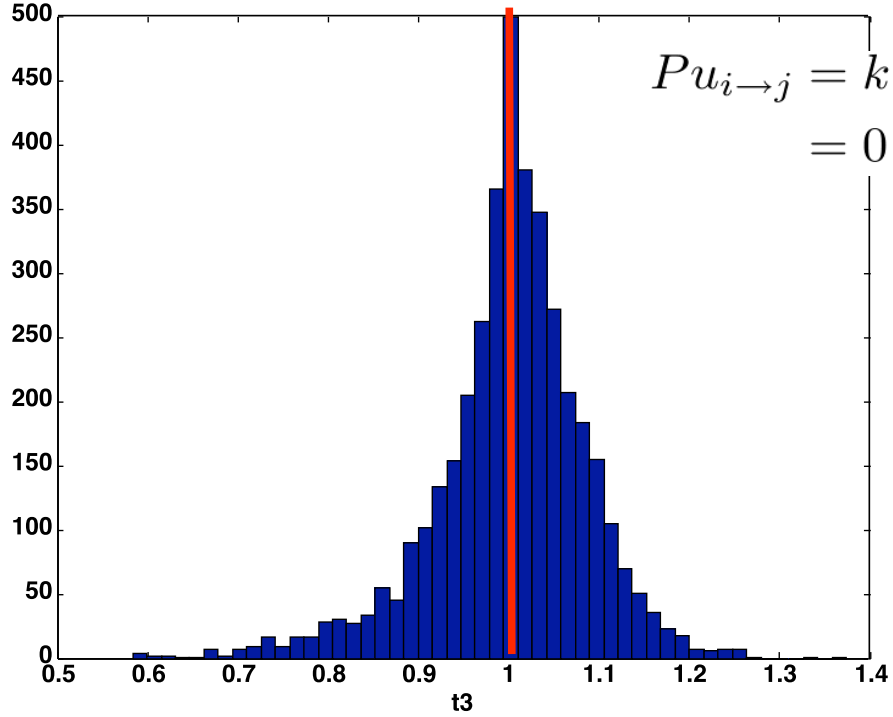


$$Pu_{i\to j} = k(m_i \cdot t_3 - m_j) \quad , m_i > t_2 \cdot m_j$$
$$= 0 \quad , m_i \leq t_2 \cdot m_j$$

In average agents are not willing (mean t2-1 ~0%) to accept a deviation from their contribution level before they decide to punish the other agent. However the standard deviation for t1 is around ~9%.

ABM Results – parameter t3

mean ~1.00 / std ~0.09

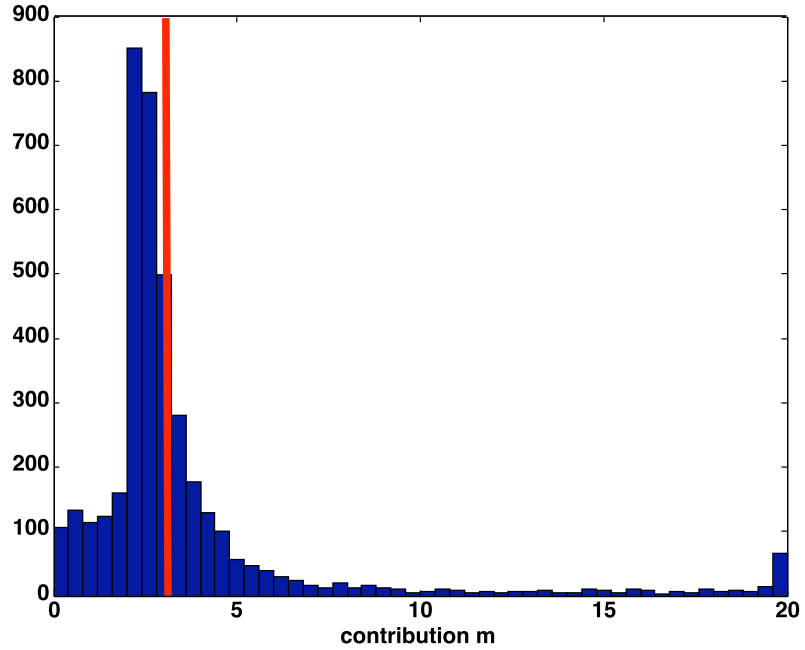


$$Pu_{i\to j} = k(m_i \cdot t_3 - m_j) \quad , m_i > t_2 \cdot m_j$$
$$= 0 \quad , m_i \leq t_2 \cdot m_j$$

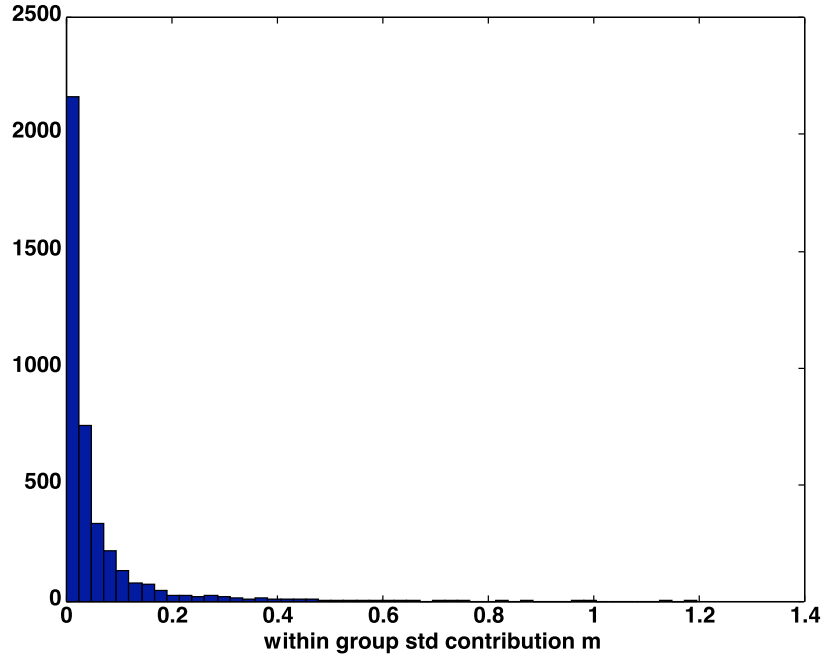
On average, agents value their contribution equivalent to the contribution of the others when they have decided to punish and calculating the amount to spent for punishing. However the standard deviation of t3 is around 9 %.

ABM Results – parameter m

mean ~3.67 / between group std ~3.63



mean of within group std ~0.05

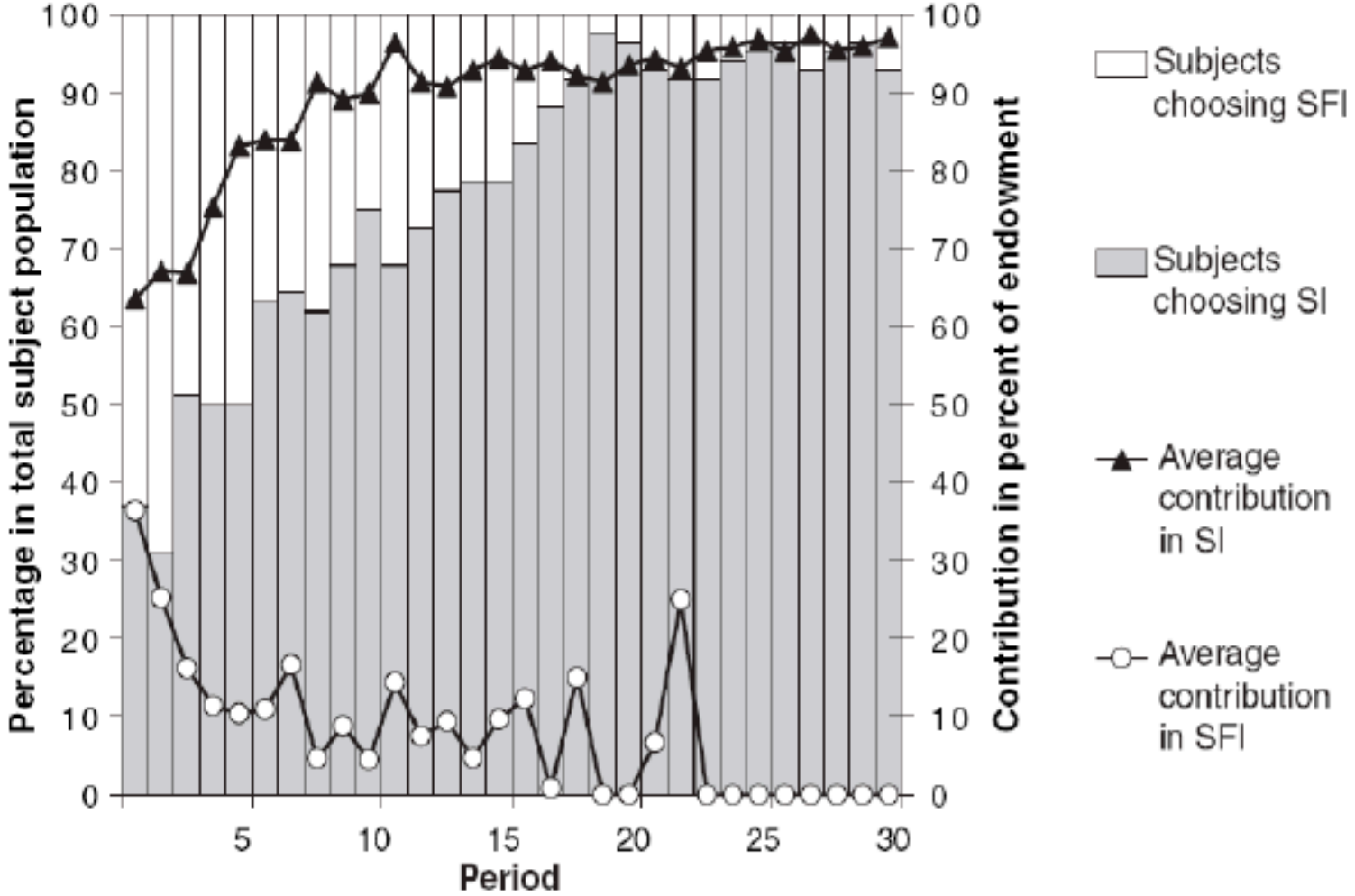


On average, agents contribute only 3.67 to the pool, but groups are quite homogeneous with an within group std. of 0.05!

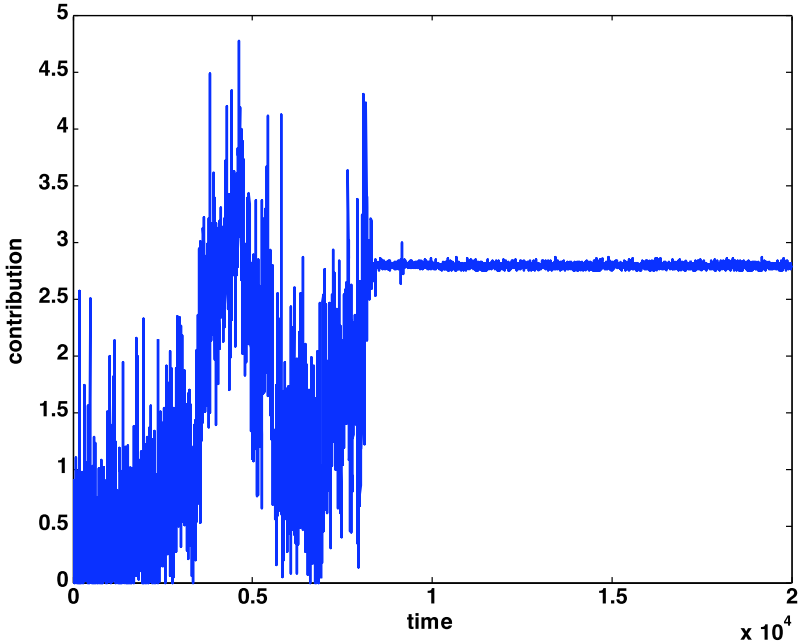
→ **Punishment seems to maintain the groups homogeneity but does not trigger cooperation!**

What happens if people can migrate between sanctioning and sanction-free institutions (Gürerk, Irlenbusch, Rockenbach Science 2006)

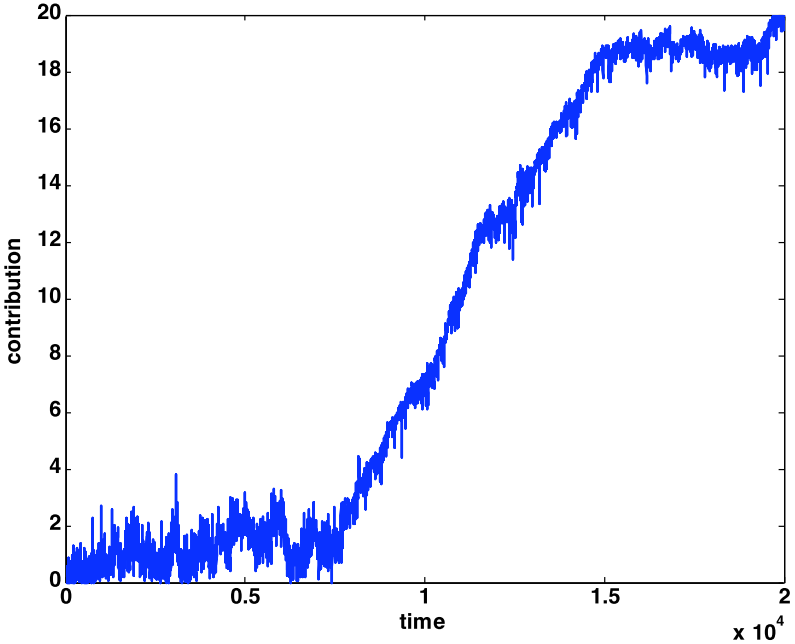
Laboratory analog of the formation of a proto state



When agents are forced to change groups randomly, cooperation emerges!

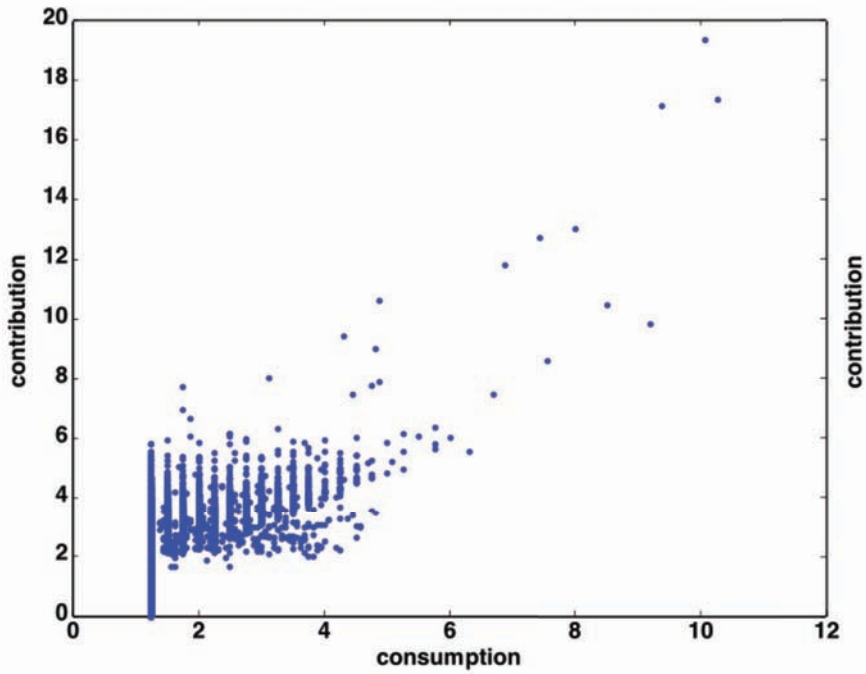


Agents do not change between groups

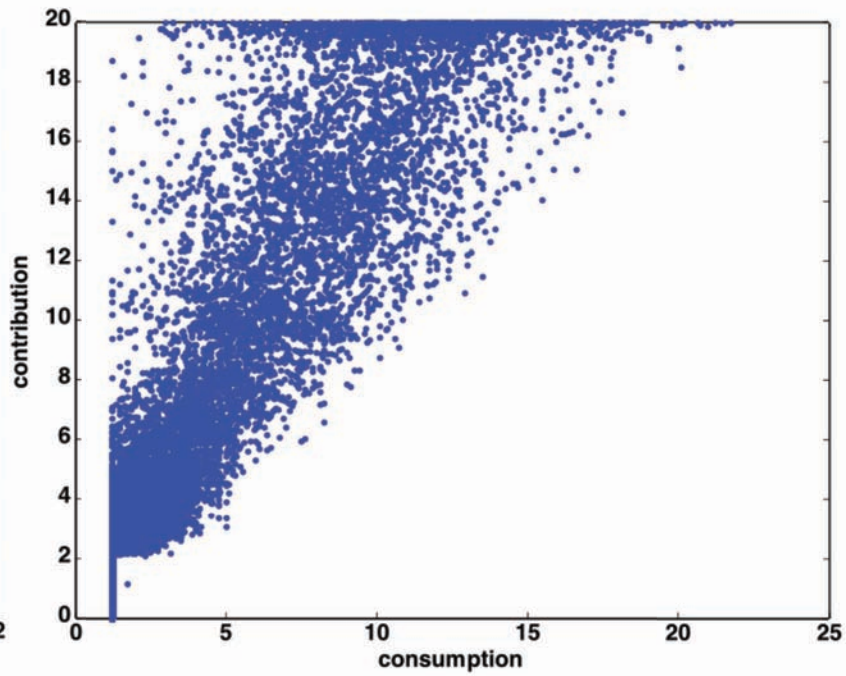


Agents change between groups

When agents are forced to change groups randomly, cooperation emerges!



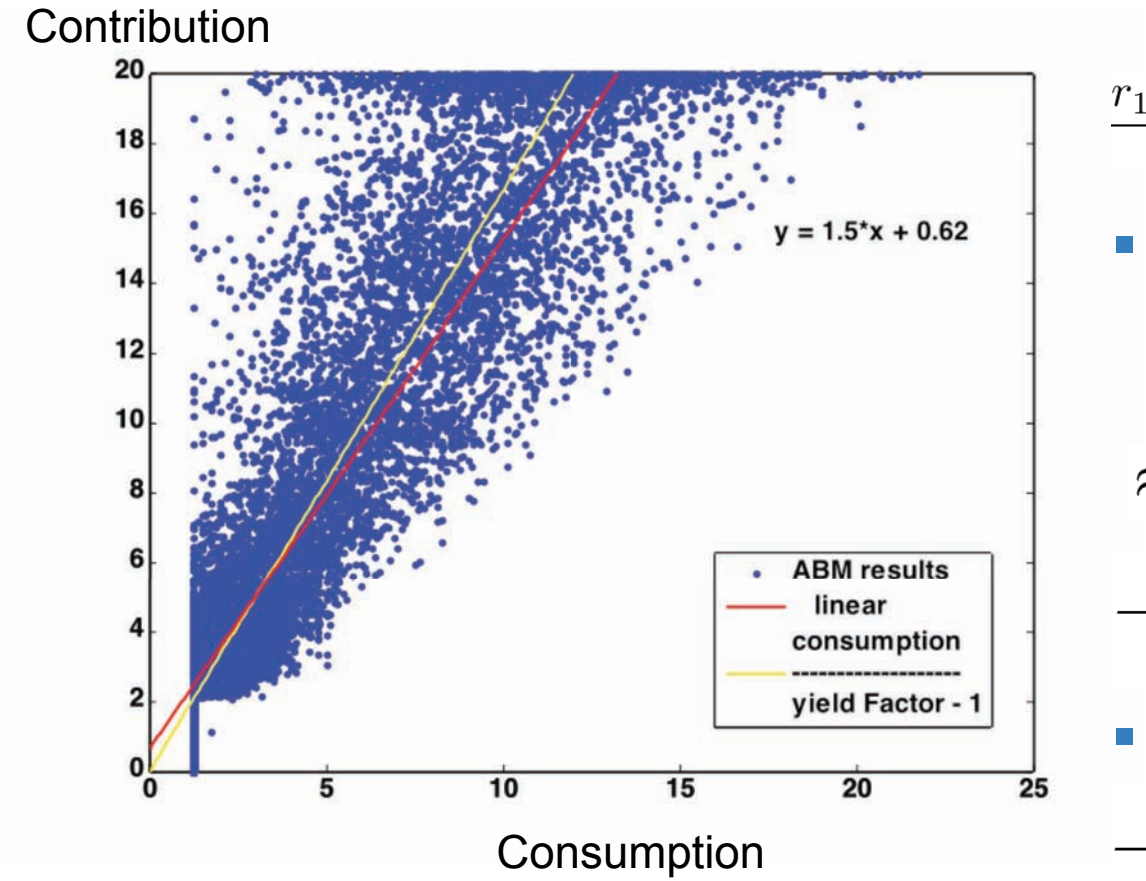
Agents do not change between groups



Agents change between groups

The role of consumption and punishment

- **Assumption:** each agent need to balance its consumption, i.e. the living expenses/costs.



$$\frac{r_1\left(E[m]\cdot(n-1)+m_i\right)}{n} - m_i = c_i$$

- if punishment is able to maintain a homogeneous group, it holds that

$$\approx m_i(r_1 - 1) = c_i$$

$$\rightarrow m_i = \frac{c_i}{r_1 - 1}$$

- With r_1=1.6, it follows:

$$\rightarrow m_i = \frac{c_i}{0.6}$$

The role of consumption and punishment

- **Assumption:** Roughly speaking each agent need to balance its consumption, i.e. the living expenses/costs.

$$\frac{r_1 \cdot E[m] \cdot (n-1) - m_i}{n} = c_i$$

- if punishment is able to maintain a homogeneous group, it holds that

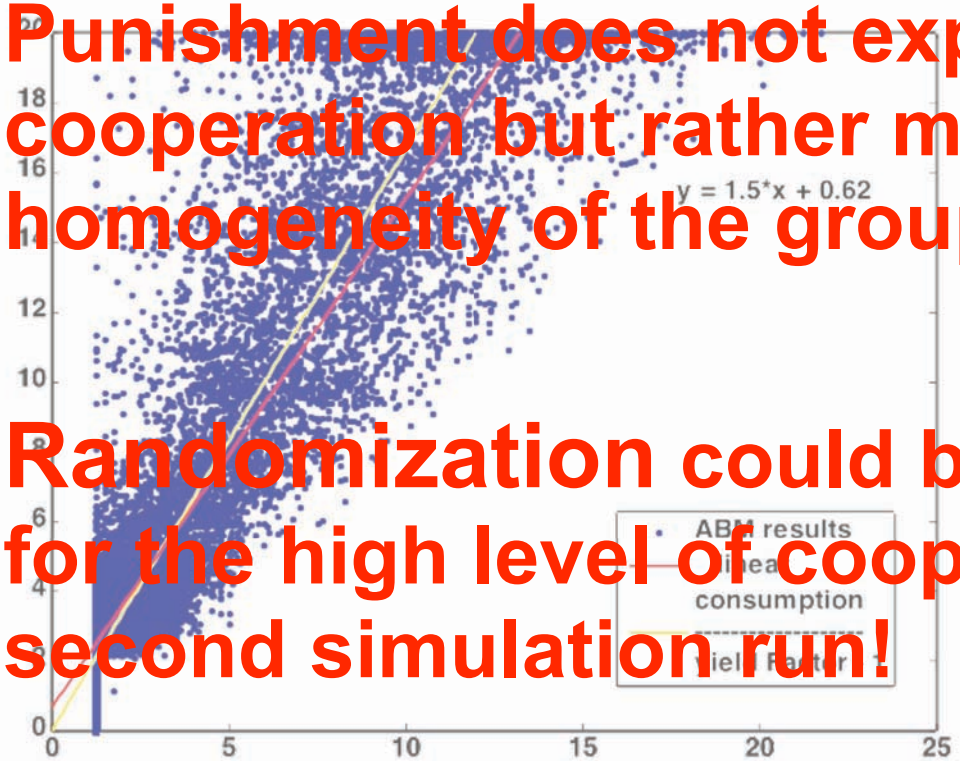
$$\approx m_i(r_1 - 1) = c_i$$

$$\rightarrow m_i = \frac{c_i}{r_1 - 1}$$

- With r_1=1.6, it follows:

$$\rightarrow m_i = \frac{c_i}{0.6}$$

- **Punishment does not explain the level of cooperation but rather maintains the homogeneity of the group!**



- **Randomization could be one reason for the high level of cooperation in the second simulation run!**

SUMMARY

•Level of punishment is determined from a robust self-centered optimization in the presence of feedbacks = attractive stable point of the dynamics of human interactions.

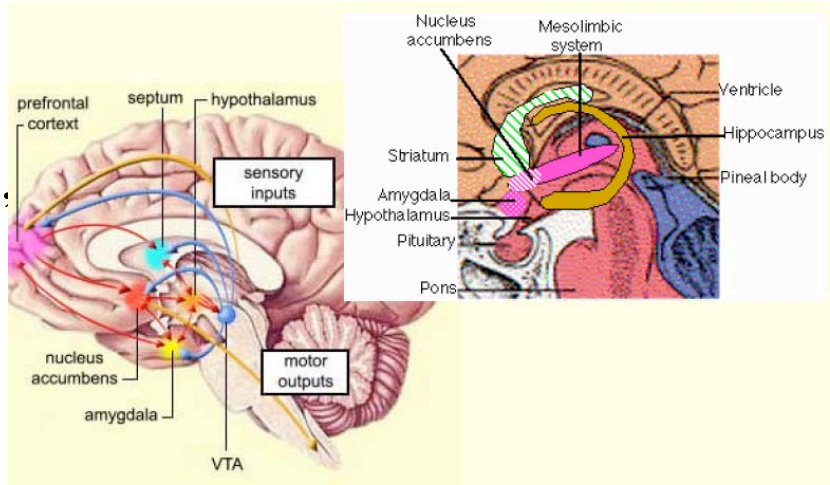
•Evolutionary feedback selection emphasizes that people have evolved to maximize their selfish gain in the presence of feedbacks.

•While the extended utility approach to strong reciprocity implies that punishment is performed as a tool to recover fairness, evolutionary feedback selection emphasizes a different driving force: punishment is not performed to recover fairness but as a tool promoting the maximum selfish gain in the presence of possible cooperation which co- evolve with it (supported by experiments)

SUMMARY

•Feedback by punishment (and probably also by reward and other mechanisms) might have been the trigger for adjusting social emotions to a level ensuring group efficiency, cascading to individual self-interest. EXPLANATION OF EMOTION-ROOTED COOPERATION

•Evolutionary feedback selection addresses human rationality in a systemic and evolutionary manner. An individual's rationality, her social emotions, her ability to understand others' desires, to transmit knowledge and social learning, influences group efficiency on a longer time scale.

•Predictions with parameter variations, limited cognitive abilities...

•Punishment ⇔ Reward ?