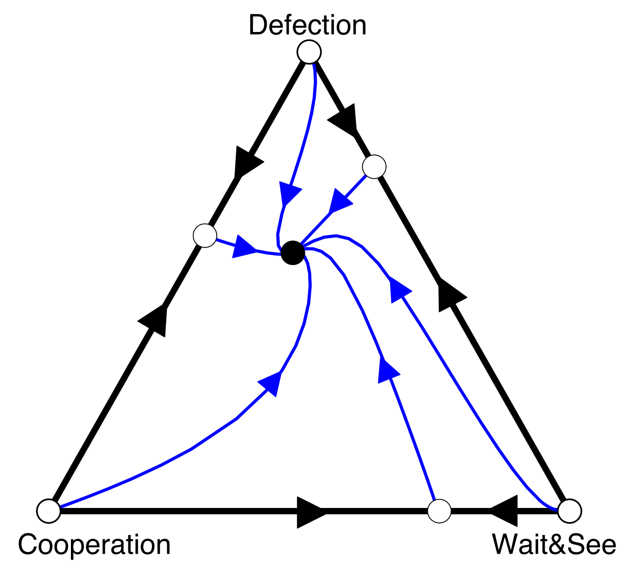
# An overview

## Yesterday's class (March 11, 2025)

- An introduction to evolutionary game theory
  (Replicator dynamics, games in finite populations)




## Today's classes (March 12, 2025)

- Evolution of cooperation & direct reciprocity
- Social norms & indirect reciprocity

## Tomorrow's class (March 13, 2025)

- Some current research: Reciprocity in complex environments

# Evolution of cooperation: Motivation

## Remark 2.1. Cooperation

In theoretical biology, cooperation is often interpreted as a costly behavior that benefits someone else.

Examples:

- Donating money to charities; providing first aid
- Acting as a referee for a journal; doing outreach
- Organising a scientific meeting (say, on decisions, games, and evolution)
- Sharing food; engaging in predator inspection; taking the lead in flocks of birds; etc.

"Proximate explanations":

- It "feels" like the right thing to do (Emotions)
- It is expected to cooperate in these situations (Social norms)
- To some extent, cooperation is mandated by law (Institutions)

Still, one might raise the question: Why would people have these particular emotions, social norms, and institutions?

NATURE VOL. 325 29 JANUARY 1987 LETTERS

**TIT FOR TAT in sticklebacks and the evolution of cooperation**

**Manfred Milinski**

Arbeitsgruppe für Verhaltensforschung, Fakultät f
Ruhr-Universität, Postfach 102148, 4630 Bochum

**Reciprocal food sharing in the vampire bat**

**Gerald S. Wilkinson**

Department of Biology, C-016, University of Ca
San Diego, La Jolla, California 92093, USA

**Matching times of leading and following suggest cooperation through direct reciprocity during V-formation flight in ibis**

Bernhard Voelkl[a,b,c,1], Steven J. Portugal[d,2], Markus Unsöld[e,f], James R. Usherwood[d], Alan M. Wilson[d], and Johannes Fritz[c,e]

# Evolution of cooperation: Prisoner's dilemma

TARGET REVIEW
**The evolution of cooperation and altruism – a general framework and a classification of models**

L. LEHMANN*† & L. KELLER*

**Five Rules for the Evolution of Cooperation**
Martin A. Nowak

**Eleven mechanisms for the evolution of cooperation**

MICHAEL A. ZAGGL*
*TUM School of Management, Technische Universität München, Arcisstr. 21, 80333, Munich, Germany*

## Remark 2.2. Cooperation in a Prisoner's dilemma

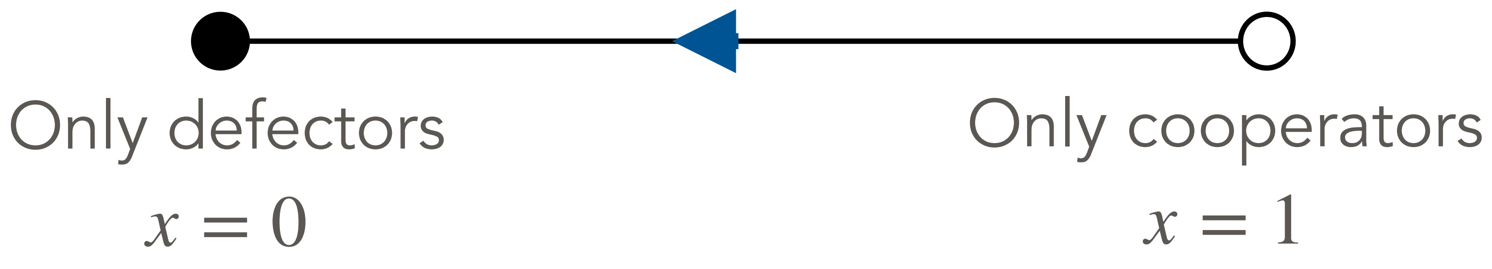
The problem of cooperation is usually illustrated with a prisoner's dilemma. Cooperation means to pay a cost *c*, for the co-player to get a payoff of *b*.

| | Cooperate | Defect |
|---|---|---|
| Cooperate | *b-c* | *-c* |
| Defect | *b* | *0* |

**Prediction based on classical game theory:**
Cooperation is a dominated strategy, hence players should not use it.

**Prediction based on evolutionary game theory:**
According to replicator dynamics, all (interior) orbits converge to Defection.

$\dot{x} = x(1-x)(-1)$




## Remark 2.3. Evolution of cooperation

So how can we explain that we actually do see quite a bit of cooperation in nature and society?

Most likely answer: Because the games people play are not fully captured by the simple prisoner's dilemma.

- We have ignored that individuals may be related to each other ( → "Kin selection")
- Evolutionary competition may also occur on the level of communities ( → "Group selection")
- Individuals may interact with each other more than once ( → "Direct reciprocity")
- Whether I cooperate with someone may affect how third parties treat me ( → "Indirect reciprocity" or "Partner choice")

# Direct reciprocity: Repeated Prisoner's dilemma (RPD)

| | C | D |
|---|---|---|
| C | b-c | -c |
| D | b | 0 |

$\xrightarrow{\delta}$

| | C | D |
|---|---|---|
| C | b-c | -c |
| D | b | 0 |

$\xrightarrow{\delta}$ …

## Definition 2.4. Repeated prisoner's dilemma

To model whether cooperation can emerge when individuals may interact more than once, we consider the repeated prisoner's dilemma.

- **Players:** The game takes place among two individuals
- **Actions:** The game takes place among multiple rounds (after each round, there is another round with probability $\delta > 0$). In each round, players can either cooperate or defect.
- **Order of moves:** In each round, individuals make their decision simultaneously.
- **Information:** Players know what happened in all previous rounds.
- **Payoffs:** Payoffs $\pi_i(t)$ in each round $t$ are the payoffs of the usual prisoner's dilemma. For the entire repeated prisoner's dilemma, we take the weighted mean:

$$\pi_i = (1-\delta)\sum_{t=1}^{\infty}\delta^t \cdot \pi_i(t)$$

[if $\delta < 1$]

$$\pi_i = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\pi_i(t)$$

[If $\delta = 1$]

## Remark 2.5. Strategies for the repeated prisoner's dilemma

In bi-matrix games, strategies were just probability distributions over the actions of the game.

In the repeated prisoner's dilemma, strategies are more complex: In each round they need to tell the player what to do, given the outcome of the previous rounds. Formally:

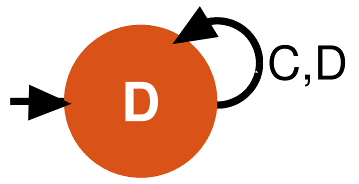
- Let $\mathbf{a}(t) = \big(a_1(t), a_2(t)\big)$ denote the outcome of round $t$, with $a_1(t), a_2(t) \in \{C, D\}$.
- A tuple $h_t = \big(\mathbf{a}(1), \mathbf{a}(2), \ldots, \mathbf{a}(t)\big)$ is a *history of the game up to round $t$*. Let $H_t$ denote the set of all such histories.
- The set $H = \cup_{t=1}^{\infty} H_t$ is the *set of all possible histories*
- A strategy $\sigma$ is now a function that takes as input the possible histories, and as output an action:

$$\sigma : H \to \{C, D\}$$

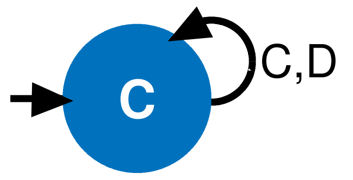
# Direct reciprocity: Repeated Prisoner's dilemma (RPD)

## Example 2.6. Some strategies for the RPD

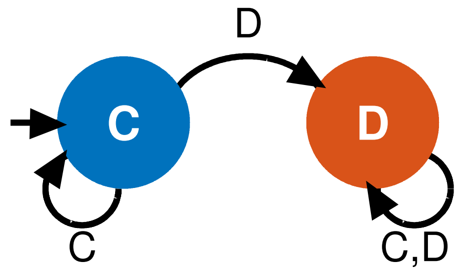
- **ALLD:** Defect in every single round, after all histories.



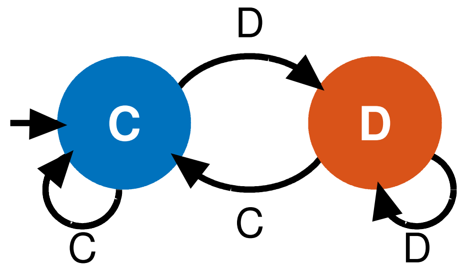

- **ALLC:** Cooperate in every single round, after all histories




- **GRIM:** Cooperate until co-player defects for the first time. Thereafter, defect forever.




- **Tit-for-Tat (TFT):** Cooperate in the first round. Thereafter, repeat whatever co-player did in previous round.




However, strategies can also become arbitrarily complicated:

- If the number of rounds played so far is prime, defect. Otherwise, cooperate if and only if the co-player cooperated twice as often as the focal player, across all rounds so far.

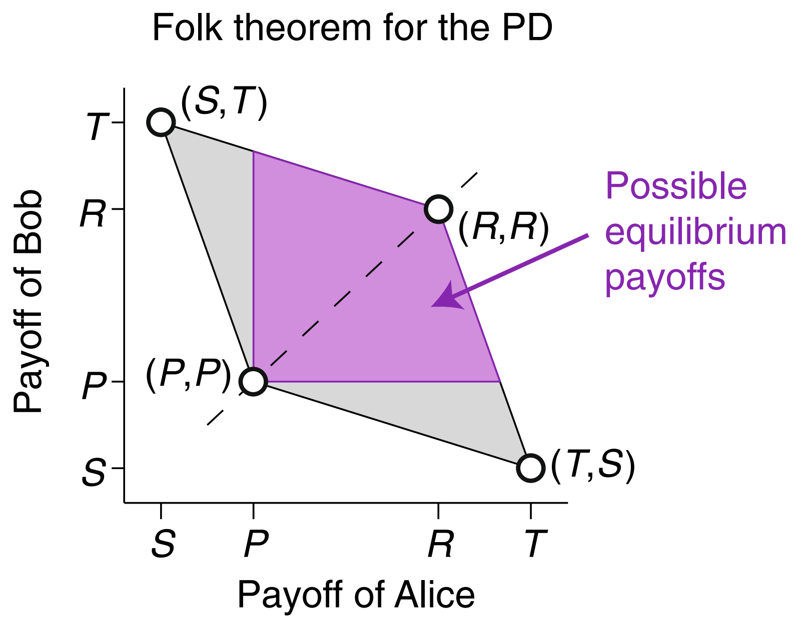
## Definition 2.7. Nash equilibrium

A strategy profile $(\sigma_1^*, \sigma_2^*)$ is a Nash equilibrium for the repeated game if

$$\pi_1(\sigma, \sigma_2^*) \leq \pi_1(\sigma_1^*, \sigma_2^*) \quad \text{and} \quad \pi_2(\sigma_1^*, \sigma) \leq \pi_2(\sigma_1^*, \sigma_2^*) \quad \text{for all } \sigma$$

# Direct reciprocity: Repeated Prisoner's dilemma (RPD)

| | C | D |
|---|---|---|
| C | $b$-$c$ | -$c$ |
| D | $b$ | 0 |




## Remark 2.8. Folk theorem of repeated games

**Question:** Suppose one knows the Nash equilibria of the one-shot (non-repeated) game. To which extent can the equilibrium payoffs $(\pi_1^*, \pi_2^*)$ of the repeated game be any different?

**Obvious requirement #1:** Even in the repeated game, equilibrium payoffs need to be *feasible* (they need to be in the convex hull of all one-shot payoffs).

**Obvious requirement #2:** Equilibrium payoffs need to be individually rational (in our case, each player needs to get at least a payoff of zero).

It turns out that these two necessary conditions are, in some sense, also sufficient:

**Folk Theorem:** Suppose $(\pi_1^*, \pi_2^*)$ is feasible and individually rational. Then for $\delta \to 1$, there is an equilibrium $(\sigma_1^*, \sigma_2^*)$ such that the corresponding payoffs in this equilibrium approach $(\pi_1^*, \pi_2^*)$.

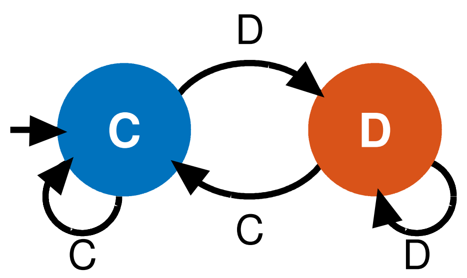
## Remark 2.9. On the Folk theorem

- Constructive proof.
- In principle, the Folk theorem is good news for our quest to understand why people may cooperate in the prisoner's dilemma. It implies that if individuals just interact sufficiently often (if $\delta$ is sufficiently large), there are equilibrium strategies according to which players cooperate in every single round.
- On the other hand, it is somewhat unsatisfactory:
  1. In some sense, there are too many equilibria ("anything goes")
  2. It remains unclear what strategy one should use in the repeated prisoner's dilemma in general.

# Direct reciprocity: Axelrod's tournament

## Remark 2.10. Axelrod & Hamilton (1981)

- Robert Axelrod invited 14 researchers to submit strategies who would be pitted against each other in a round-robin tournament.
- All participants knew the possible payoffs in each round, but they did not know how many rounds each game would take (this number was 200).
- Some strategies were quite sophisticated.
- The simplest (shortest) submitted strategy won the tournament: Tit-for-Tat.




**The Evolution of Cooperation**

Robert Axelrod and William D. Hamilton

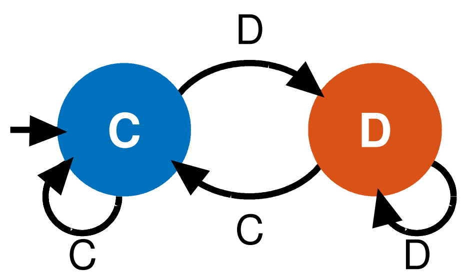
# Direct reciprocity: Axelrod’s tournament

**The Evolution of Cooperation**

Robert Axelrod and William D. Hamilton

## Remark 2.10. Axelrod & Hamilton (1981)

- Robert Axelrod invited 14 researchers to submit strategies who would be pitted against each other in a round-robin tournament.
- All participants knew the possible payoffs in each round, but they did not know how many rounds each game would take (this number was 200).
- Some strategies were quite sophisticated.
- The simplest (shortest) submitted strategy won the tournament: Tit-for-Tat.




| | Round 1 | Round 2 | Round 3 | Round 4 | … |
|---|---|---|---|---|---|
| TFT | C | D | D | D | ⋯ |
| ALLD | D | D | D | D | ⋯ |
| Payoffs | (-c,b) | (0,0) | (0,0) | (0,0) | … |

Average payoffs:

$$\pi_{TFT} = -(1-\delta)c$$

$$\pi_{ALLD} = (1-\delta)b$$

| | Round 1 | Round 2 | Round 3 | Round 4 | … |
|---|---|---|---|---|---|
| TFT | C | C | C | C | ⋯ |
| TFT | C | C | C | C | ⋯ |
| Payoffs | (b-c,b-c) | (b-c,b-c) | (b-c,b-c) | (b-c,b-c) | … |

$$\pi_{TFT} = b - c$$

Tit-for-Tat is a best response against itself if $b - c > (1-\delta)b$, or if $\delta > c/b$.

NATURE VOL. 325 29 JANUARY 1987 LETTERS

**TIT FOR TAT in sticklebacks and the evolution of cooperation**

**Manfred Milinski**

Arbeitsgruppe für Verhaltensforschung, Fakultät für Biologie, Ruhr-Universität, Postfach 102148, 4630 Bochum 1, FRG

**Reciprocal food sharing in the vampire bat**

**Gerald S. Wilkinson**

Department of Biology, C-016, Uni... San Diego, La Jolla, California 9209...

# Direct reciprocity: Evolution of cooperation

## Remark 2.11. On Axelrod & Hamilton (1981)

- The paper by Axelrod & Hamilton has been hugely influential for the field.
- However, Tit-for-Tat is sensitive to errors.
- More crucially, the results of such a tournament very much depend on the participating strategies.
- It would be nice to have results for an unbiased strategy space. This strategy space should be sufficiently large to be interesting, but sufficiently small to be manageable.

| | Rd. 1 | Rd. 2 | Rd. 3 | Rd. 4 | Rd. 5 | … |
|---|---|---|---|---|---|---|
| TFT | C | C | **D*** | C | D | … |
| TFT | C | C | C | D | C | … |

## Remark 2.12. Memory-1 strategies

- Suppose players only condition their behavior on the outcome of the previous round.
- The corresponding memory-1 strategies are vectors $\mathbf{p} = (p_0, p_{CC}, p_{CD}, p_{DC}, p_{DD}) \in [0,1]^5$
- Examples: ALLD = (0,0,0,0,0); TFT = (1,1,0,1,0).
- Major advantage: Games among memory-1 players can be represented as a Markov chain. The states of this Markov chain are the four possible outcomes of a given round, (C,C), (C,D), (D,C), (D,D).

$$M = \begin{pmatrix} p_{CC}q_{CC} & p_{CC}(1-q_{CC}) & (1-p_{CC})q_{CC} & (1-p_{CC})(1-q_{CC}) \\ p_{CD}q_{DC} & p_{CD}(1-q_{DC}) & (1-p_{CD})q_{DC} & (1-p_{CD})(1-q_{DC}) \\ p_{DC}q_{CD} & p_{DC}(1-q_{CD}) & (1-p_{DC})q_{CD} & (1-p_{DC})(1-q_{CD}) \\ p_{DD}q_{DD} & p_{DD}(1-q_{DD}) & (1-p_{DD})q_{DD} & (1-p_{DD})(1-q_{DD}) \end{pmatrix}$$

- In particular, for any pair of memory-1 strategies $\mathbf{p}$ and $\mathbf{q}$, one can compute the players' payoffs $\pi_1(\mathbf{p},\mathbf{q})$ and $\pi_2(\mathbf{p},\mathbf{q})$.

# Direct reciprocity: Evolution among Memory-1 strategies

## Remark 2.13. Nowak and Sigmund (1993)

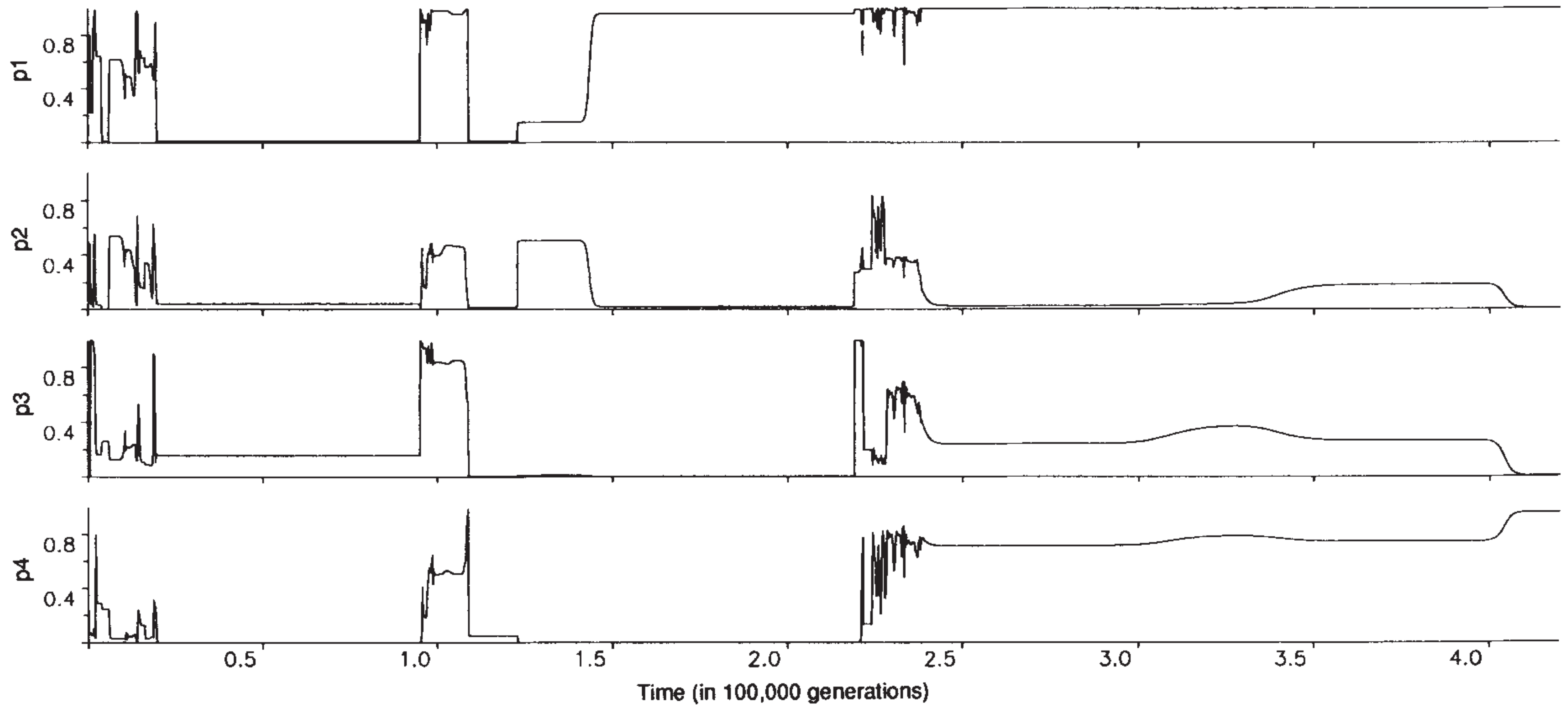
- Approach: Let evolution determine an "optimal" strategy, by running individual-based simulations.
- Consider a finite population of size $N$
- Initially, all players use random memory-1 strategies.
- Players obtain payoffs by playing the repeated prisoner's dilemma with all other population members.
- Afterwards, players reproduce proportional to their fitness. In addition, mutations may introduce new strategies.
- Result: Most simulations lead to Win-Stay Lose-Shift. This strategy is robust against errors, it can exploit unconditional cooperators, and it is a Nash equilibrium if $b>2c$.

  [For memory-k strategies, one can find generalised versions of this strategy that are stable for $b > \frac{k+1}{k}c$]

**A strategy of win–stay, lose–shift that outperforms tit-for-tat in the Prisoner's Dilemma game**

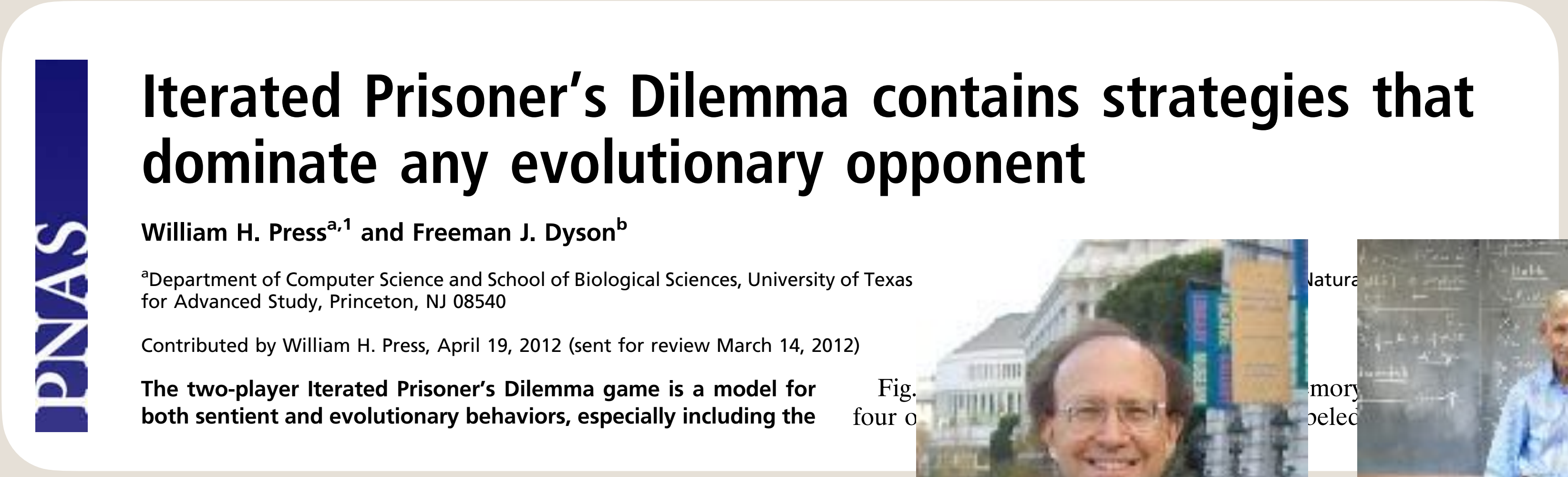
**Martin Nowak* & Karl Sigmund†**

*Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK
†Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A-1090 Vienna, Austria




**Win-stay Lose-shift WSLS = (1,1,0,0,1)**
Round 1: Cooperate
All other rounds:
- After obtaining b or b-c, do what you did last round
- Otherwise do the opposite of what you did.

# Direct reciprocity: Direct reciprocity and extortion



PNAS

Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent

William H. Press[a,1] and Freeman J. Dyson[b]

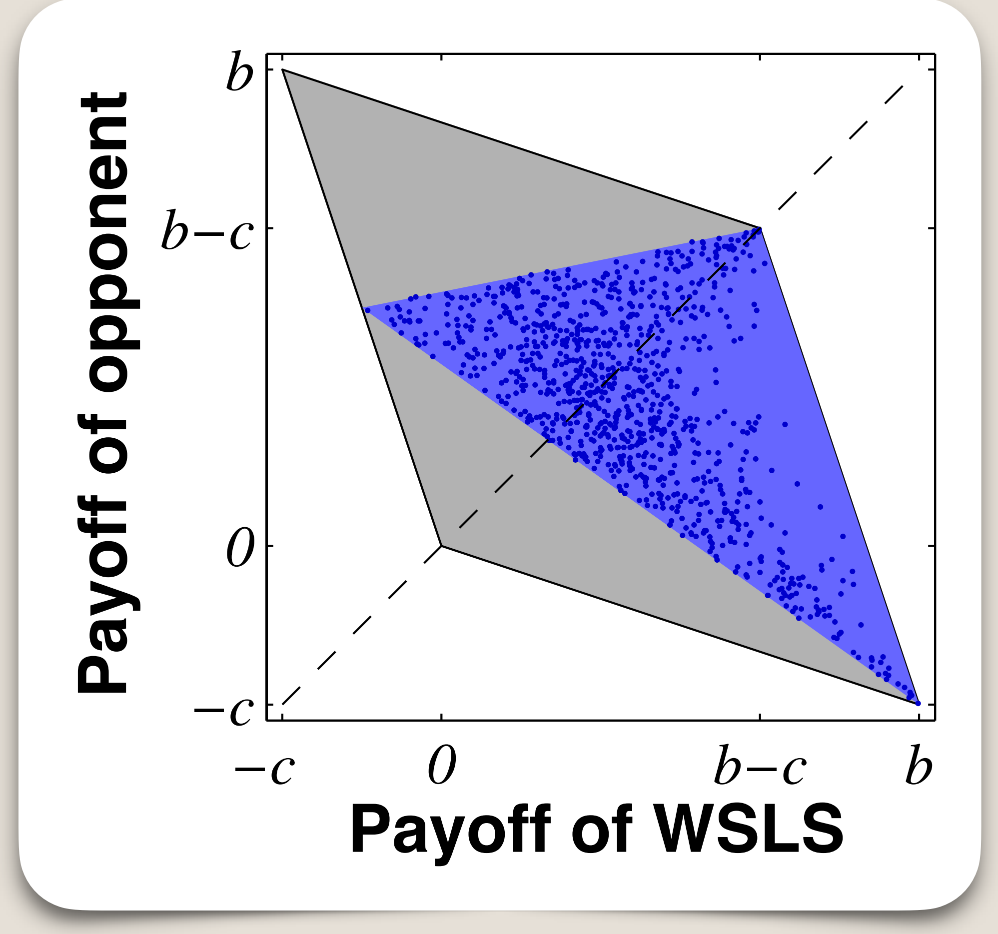
[a]Department of Computer Science and School of Biological Sciences, University of Texas
for Advanced Study, Princeton, NJ 08540

Contributed by William H. Press, April 19, 2012 (sent for review March 14, 2012)

The two-player Iterated Prisoner's Dilemma game is a model for both sentient and evolutionary behaviors, especially including the

## Remark 2.14. Press & Dyson (2012)

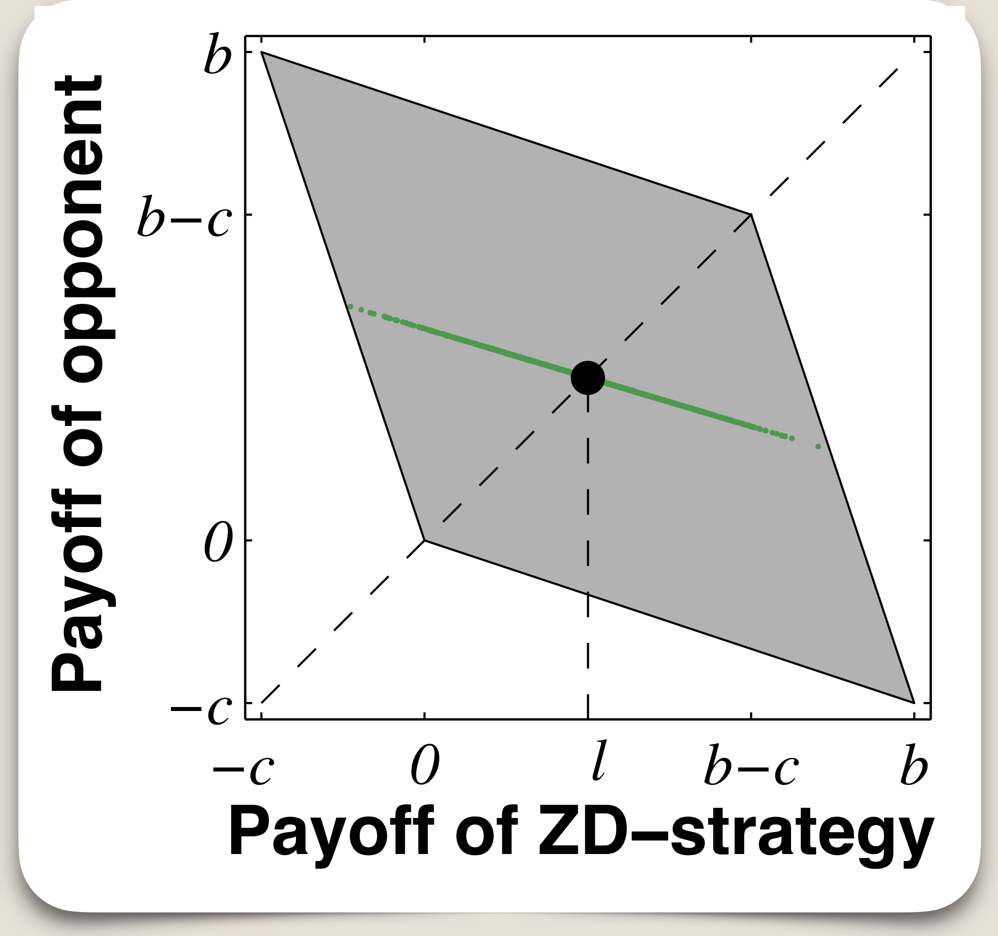
- By the 2010s, the field largely believed that the repeated prisoner's dilemma almost naturally leads to the evolution of cooperation.
- However, then William Press and Freeman Dyson derived quite a surprising result.



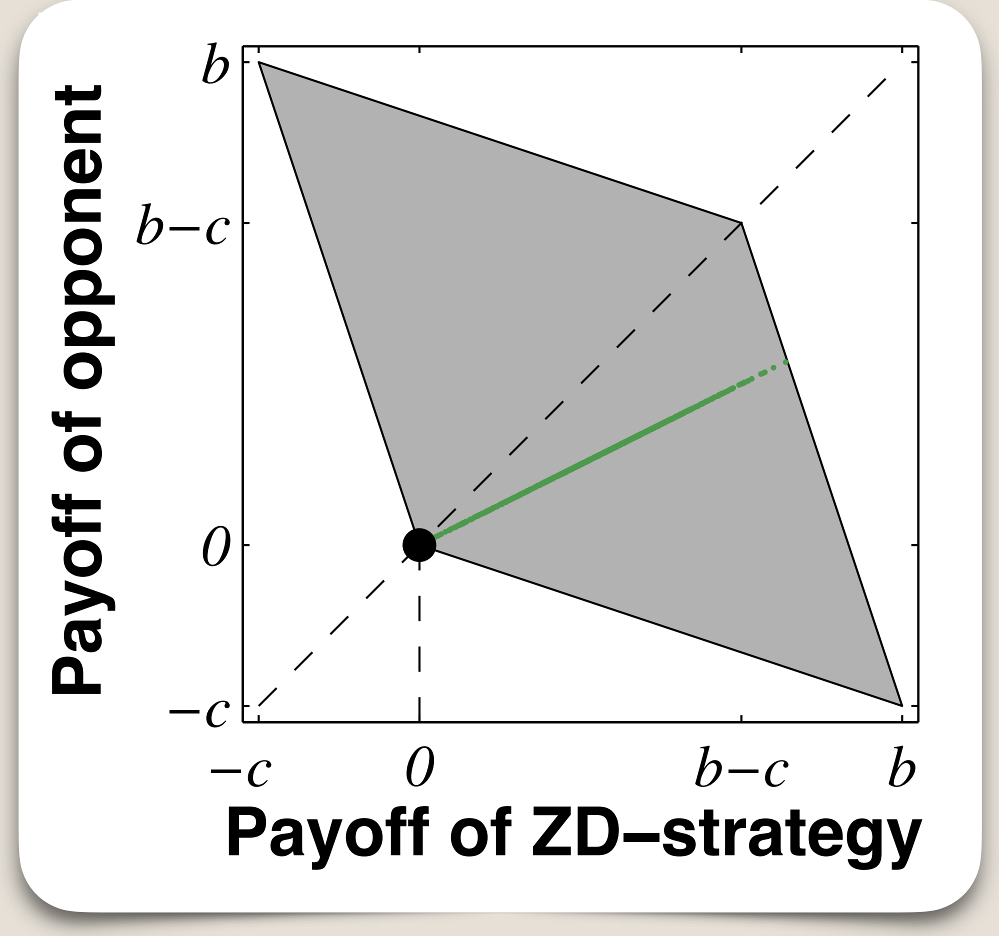

**Theorem:** Consider an infinitely repeated prisoner's dilemma with $\delta = 1$. Suppose there are constants $\alpha, \beta, \gamma$ such that player 1 applies a memory-1 strategy $\mathbf{p}$ that satisfies

$$
\begin{aligned}
p_{CC} &= (\alpha+\beta)(b-c) + \gamma + 1 \\
p_{CD} &= -\alpha c + \beta b + \gamma + 1 \\
p_{DC} &= \alpha b - \beta c + \gamma \\
p_{DD} &= \gamma
\end{aligned}
$$

Then, irrespective of player 2's strategy, payoffs satisfy $\alpha\pi_1 + \beta\pi_2 + \gamma = 0$. Such a strategy $\mathbf{p}$ is called a zero-determinant (ZD) strategy.




*Extortionate ZD strategy*

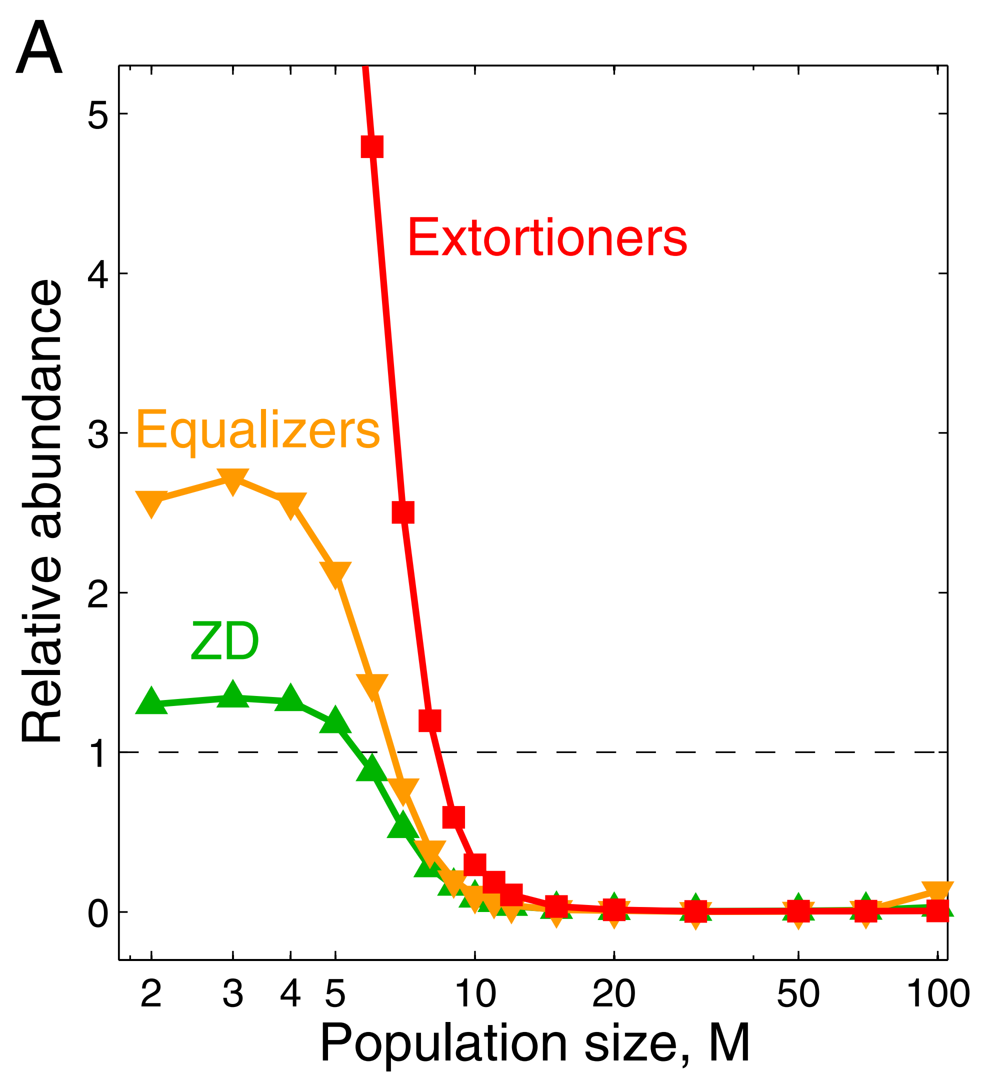


# Direct reciprocity: Evolution of extortion

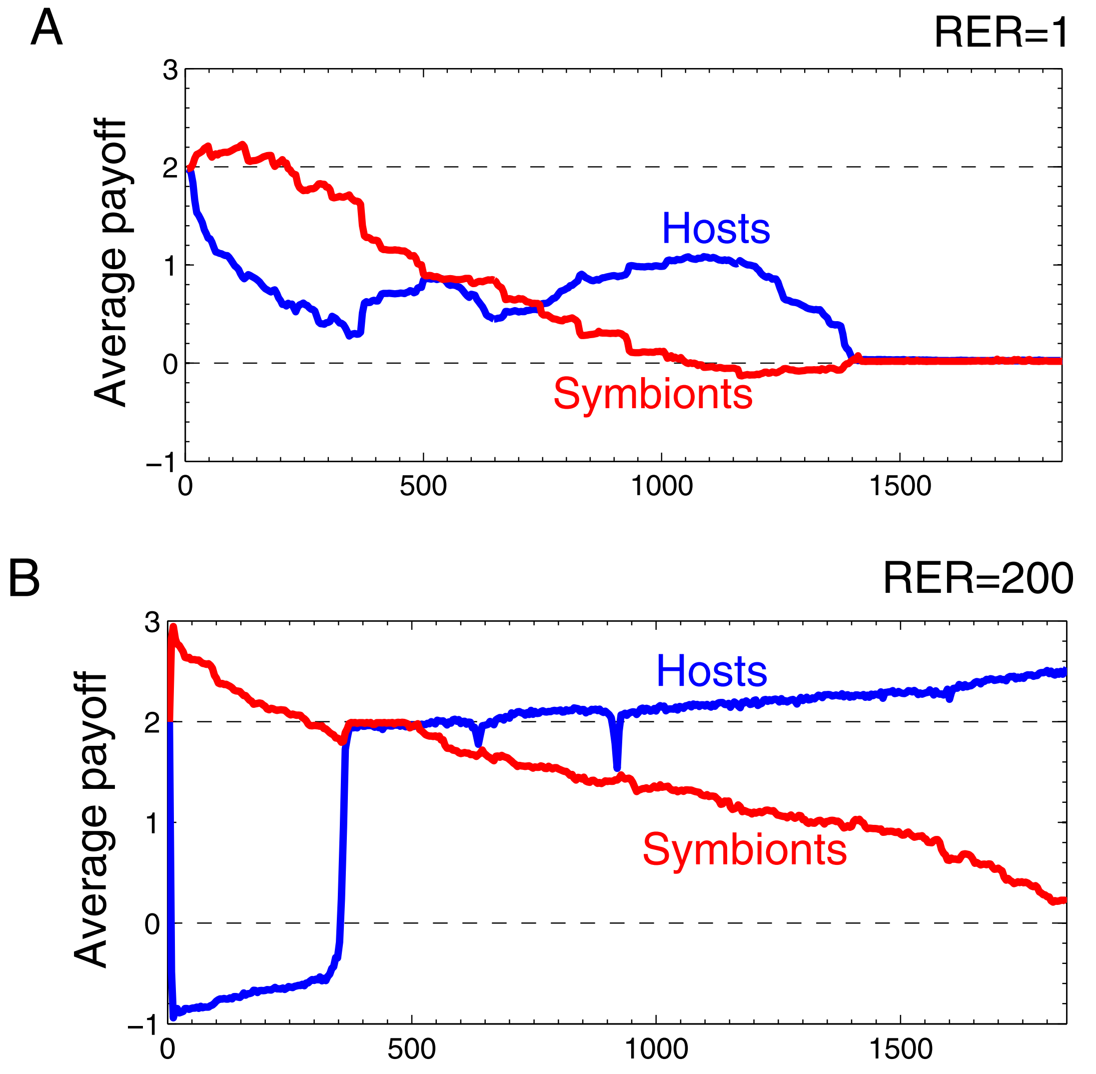
## Remark 2.15. Can extortion evolve?

- One can repeat the individual-based simulations of Nowak and Sigmund to check for which parameters we would observe the evolution of extortion.
- To this end, one can record how often players would choose a strategy in the neighbourhood of extortionate strategies.
- The paper considers two evolutionary scenarios: evolution within one population, and evolution among two co-evolving populations (who may evolve at different rates).
- Result: Extortion only evolves when populations are small, or when two populations interact and one population evolves at a much slower rate ("Red king effect")

### Evolution within one population




### Evolution with two populations




**Evolution of extortion in Iterated Prisoner's Dilemma games**

Christian Hilbe[a], Martin A. Nowak[b], and Karl Sigmund[c,d,1]

# Direct reciprocity: Extortion in the lab

Extortion subdues human players but is finally punished in the prisoner's dilemma

Christian Hilbe[1,2], Torsten Röhl[1] & Manfred Milinski[3]

## Remark 2.16. How do humans react to extortion?

- Another interesting question: Do humans use extortionate strategies?
- However, this question is difficult to test: when you let participants interact in the repeated prisoner's dilemma, you can only observe what they do (their actions), but not how their general rule to choose actions (their strategies).
- Instead one could ask: How do humans react to extortionate strategies? Is it profitable to use extortion against human participants?

- Setup of the experiment:
  - 60 participants played for 60 rounds

| | **Cooperate** | **Defect** |
|---|---|---|
| **Cooperate** | 0.30 Euro | 0.00 Euro |
| **Defect** | 0.50 Euro | 0.10 Euro |

  - Participants played against a computer program that implemented an extortionate or a generous strategy

- Predictions:
  1. *Computer vs Humans:*
     In the treatment with extortion, the computer should get a higher payoff than the participants. In the treatment with generosity, the human participants should get the higher payoff.
  2. *Dynamics of cooperation:*
     In all treatments, the best response for humans is to cooperate in every round. Hence we would expect a general trend towards cooperation in all treatments.
  3. *Extortion vs generosity:*
     If humans indeed learn to cooperate in all treatments, the extortionate program should get higher payoffs than the generous program.

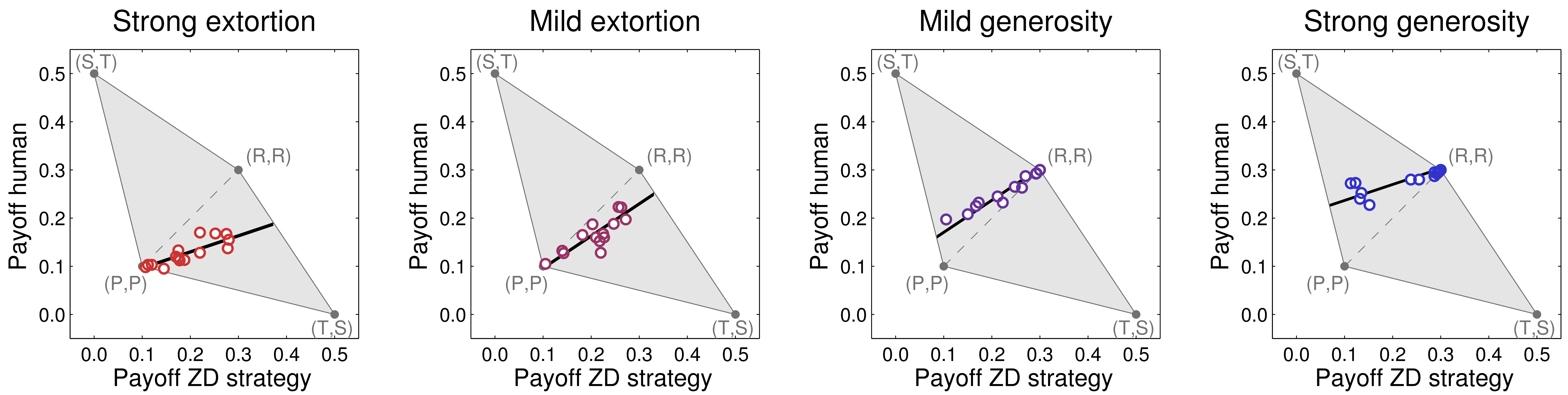
# Direct reciprocity: Extortion in the lab

Extortion subdues human players but is finally punished in the prisoner's dilemma

Christian Hilbe[1,2], Torsten Röhl[1] & Manfred Milinski[3]

## Remark 2.16. Human reactions to extortion (continued)

Prediction 1 (*Computer vs Humans*)
In the treatment with extortion, the computer should get a higher payoff than the participants. In the treatment with generosity, the human participants should get the higher payoff.



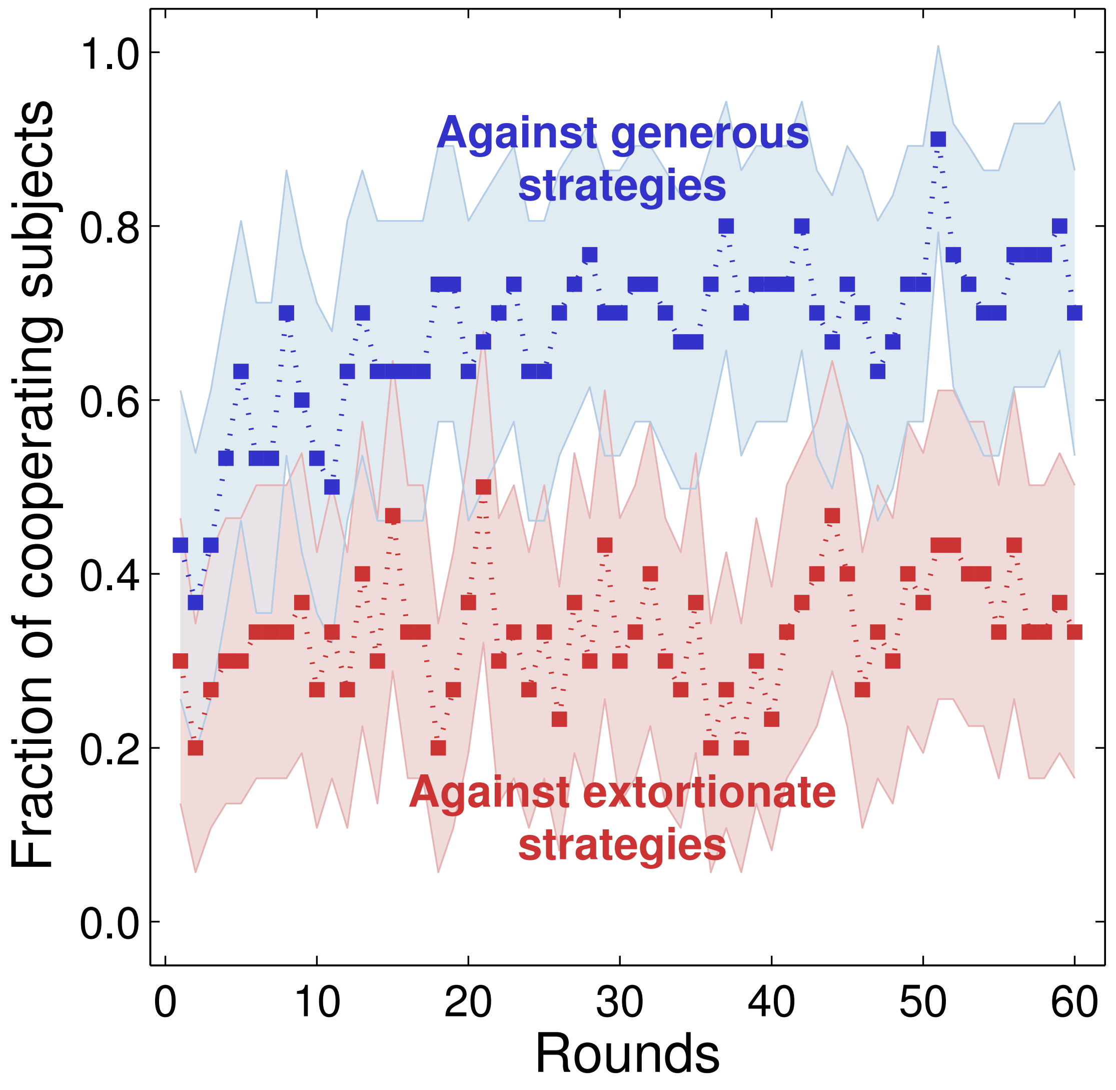
# Direct reciprocity: Extortion in the lab

Extortion subdues human players but is finally punished in the prisoner's dilemma

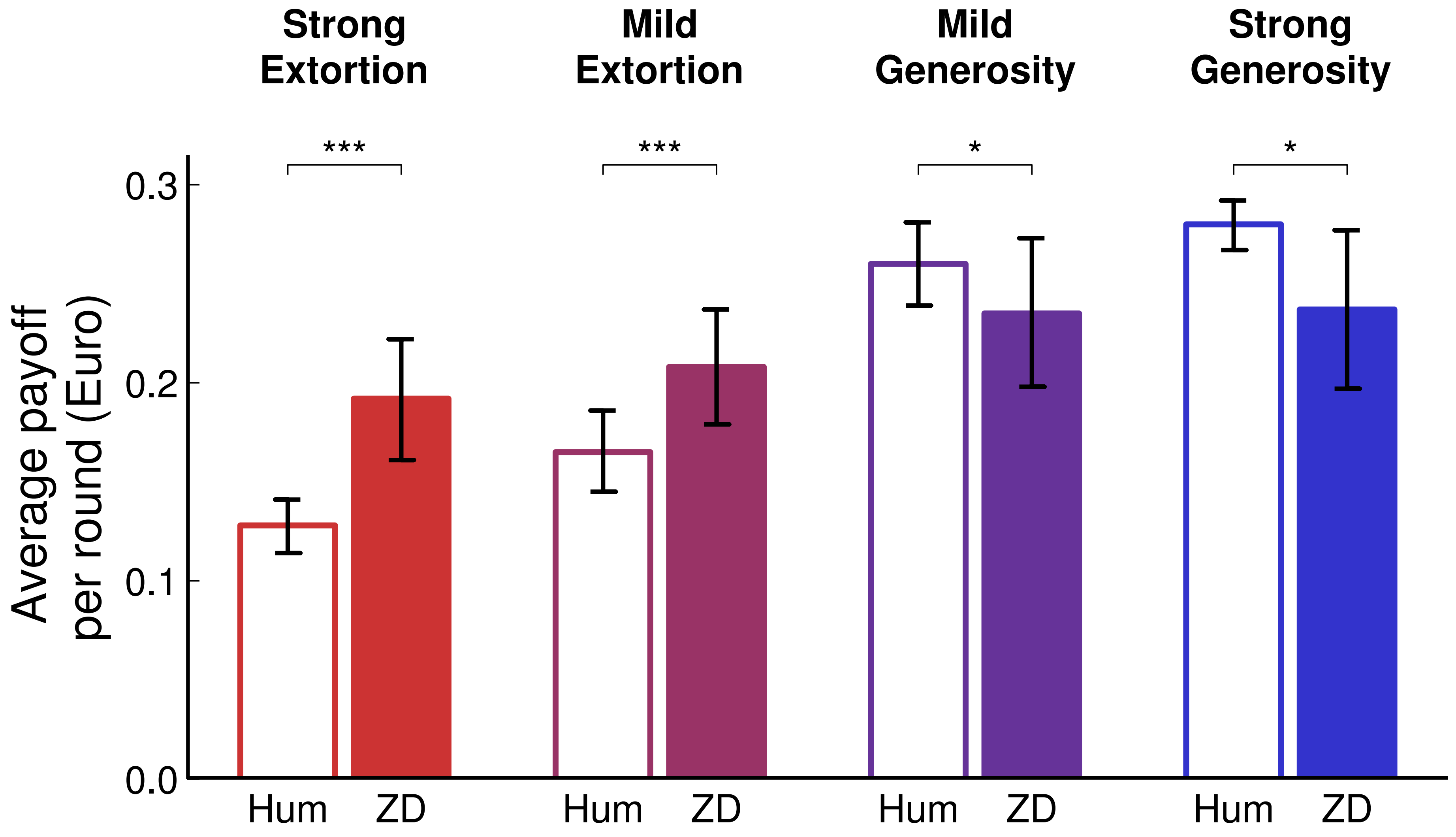
Christian Hilbe[1,2], Torsten Röhl[1] & Manfred Milinski[3]

## Remark 2.16. Human reactions to extortion (continued)

Prediction 2 (*Dynamics of cooperation*)
In all treatments, the best response for humans is to cooperate in every round. Hence we would expect a general trend towards cooperation in all treatments.




Prediction 3 (*Extortion vs generosity*)
In all treatments, the best response for humans is to cooperate in every round. Hence we would expect a general trend towards cooperation in all treatments.



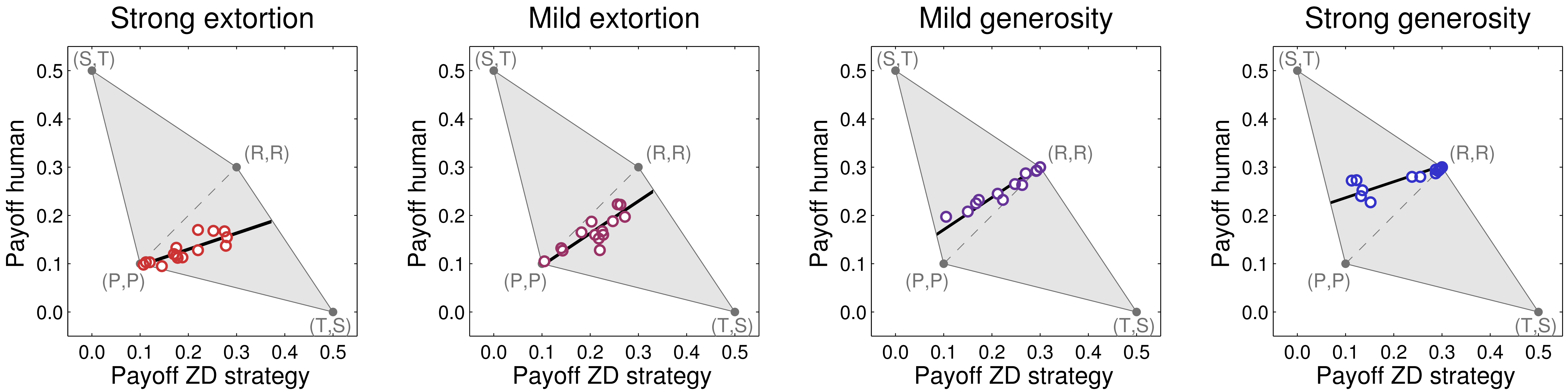
# Direct reciprocity: Extortion in the lab

Extortion subdues human players but is finally punished in the prisoner's dilemma

Christian Hilbe[1,2], Torsten Röhl[1] & Manfred Milinski[3]

## Remark 2.16. Human reactions to extortion (continued)

*Interpretation:*




- A possible explanation for these patterns is that human participants had a strong preference for fair outcomes. In the generosity treatments, payoff-maximisation and fairness are aligned. In the extortion treatments, they are mis-aligned.
- In line with this interpretation, the effect vanishes if human participants receive information about the nature of their opponent before the experiment.

# Summary

## Some things you should have learned today:

1. We have used our techniques (Defining a game, characterizing Nash equilibria, evolutionary dynamics) to explore why people may cooperate.

2. One mechanism for cooperation is direct reciprocity. It is based on the intuition that cooperation can in fact be a profitable strategy when people interact repeatedly (no conscious decision-making required!)

3. Extortionate (zero-determinant) strategies have interesting mathematical properties, but against humans they don't pay.