Incentives and Opportunism: From The Carrot to the Stick

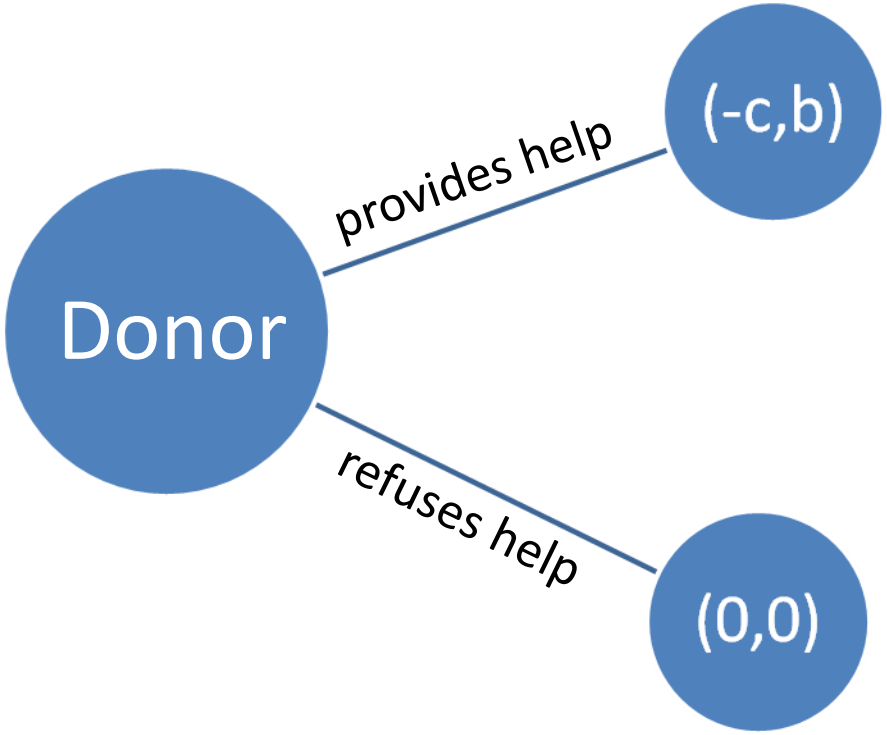
C. Hilbe & K. Sigmund, Proc. R. Soc. B (2010) 277, 2427–2433

TECT Final Conference, Budapest

From the Carrot to the Stick

Introduction Model & Results Discussion

- Helping Game



Costs of helping: -c

Benefit of help: b

Social optimum (for b>c):
Providing help

Individual optimum
Refusing help

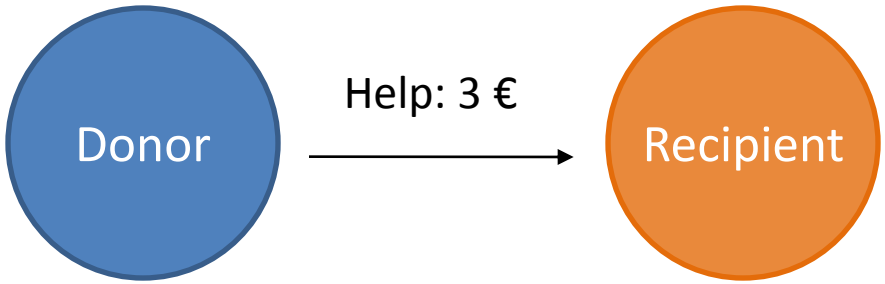
From the Carrot to the Stick

Experimental Results (Dictator Game)

	Donor	Recipient
Show-up fee	5 €	5 €
Dictator endowment	10 €	--
Total payoff	15 €	5 €

From the Carrot to the Stick

Experimental Results (Dictator Game)



Show-up fee	5 €	5 €
Dictator endowment	10 €	--
Transfer	- 3 €	3 €
Total payoff	12 €	8 €

From the Carrot to the Stick

Experimental Results (Dictator Game)

• Usually, more than 50 % of the donors transfer a positive amount (1-5 €)

But:

• 97 % of the donors refuse any help, if they have earned the 10 € in a quiz *(Cherry et al., Am. Econ. Rev. 92, 2002)*

•More than 80 % of the donors are also willing to take up to 3 € from the recipient's show-up fee *(N. Bardsley, Exp. Econ. 11, 2008)*

From the Carrot to the Stick

Conclusion

- Neither hard-nosed game theory nor pure altruism can explain the data

- Additional incentives for cooperation are needed, e.g.
  - Punishment of selfish behavior
  - Rewarding generosity

From the Carrot to the Stick

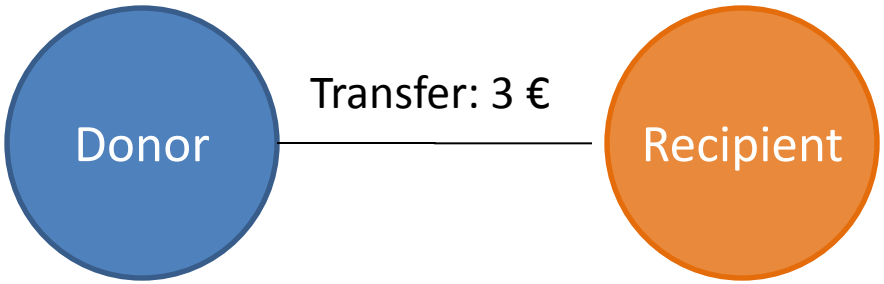
Examples: Punishment in animal societies
(T.H. Clutton-Brock & G.A. Parker, Nature, 1995)

- Rhesus macaques that find sources of preferred food and do not give food calls are more likely to be the target of aggression
- Chimpanzees form supportive coalitions to gain access to resources and attack those allies that fail to support them in competitive interactions with third parties
- In *Polistes* wasps, queens are regularly aggressive to inactive workers.

Conclusion:
Punishment is 1) used to ensure cooperation and 2) is effective in doing so

From the Carrot to the Stick

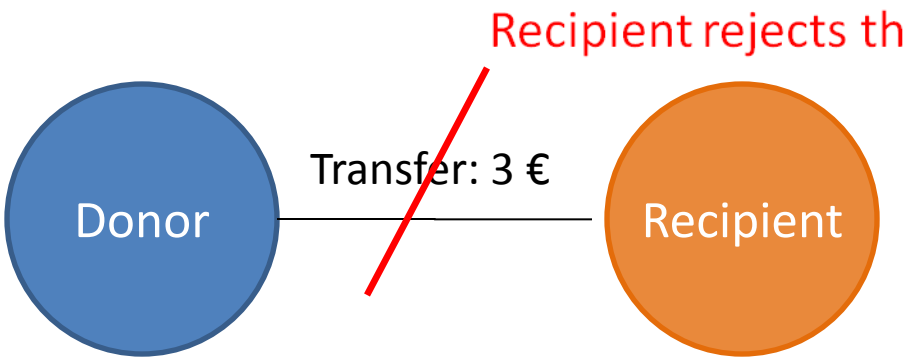
Experimental Results (Ultimatum Game)



Show-up fee	5 €	5 €
Dictator endowment	10 €	--
Transfer	- 3 €	3 €
Total payoff	12 €	8 €

From the Carrot to the Stick

Experimental Results (Ultimatum Game)



Show-up fee	5 €	5 €
Dictator endowment	~~10 €~~	~~--~~
Transfer	~~- 3 €~~	~~3 €~~
Total payoff	~~12 €~~	~~8 €~~
	5 €	**5 €**

From the Carrot to the Stick

Experimental Results (Ultimatum Game)

- Typical offers are between 3-5 €
- Lower offers are usually rejected by the recipients

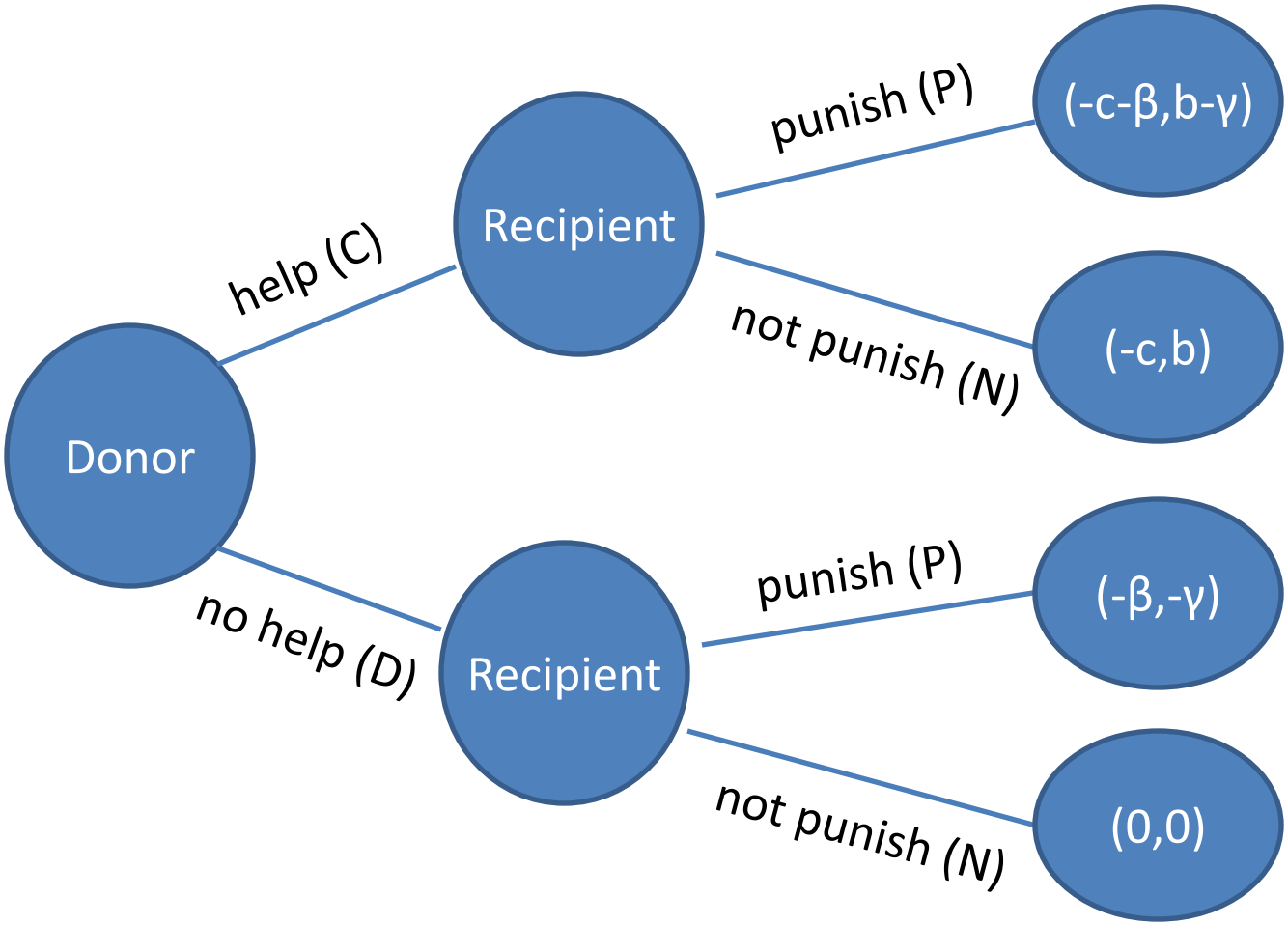
Conclusions

- Individuals are willing to punish selfish behavior even if punishment is costly
- Donors anticipate this and offer higher amounts

From the Carrot to the Stick

Introduction Model & Results Discussion

- Helping Game with Punishment

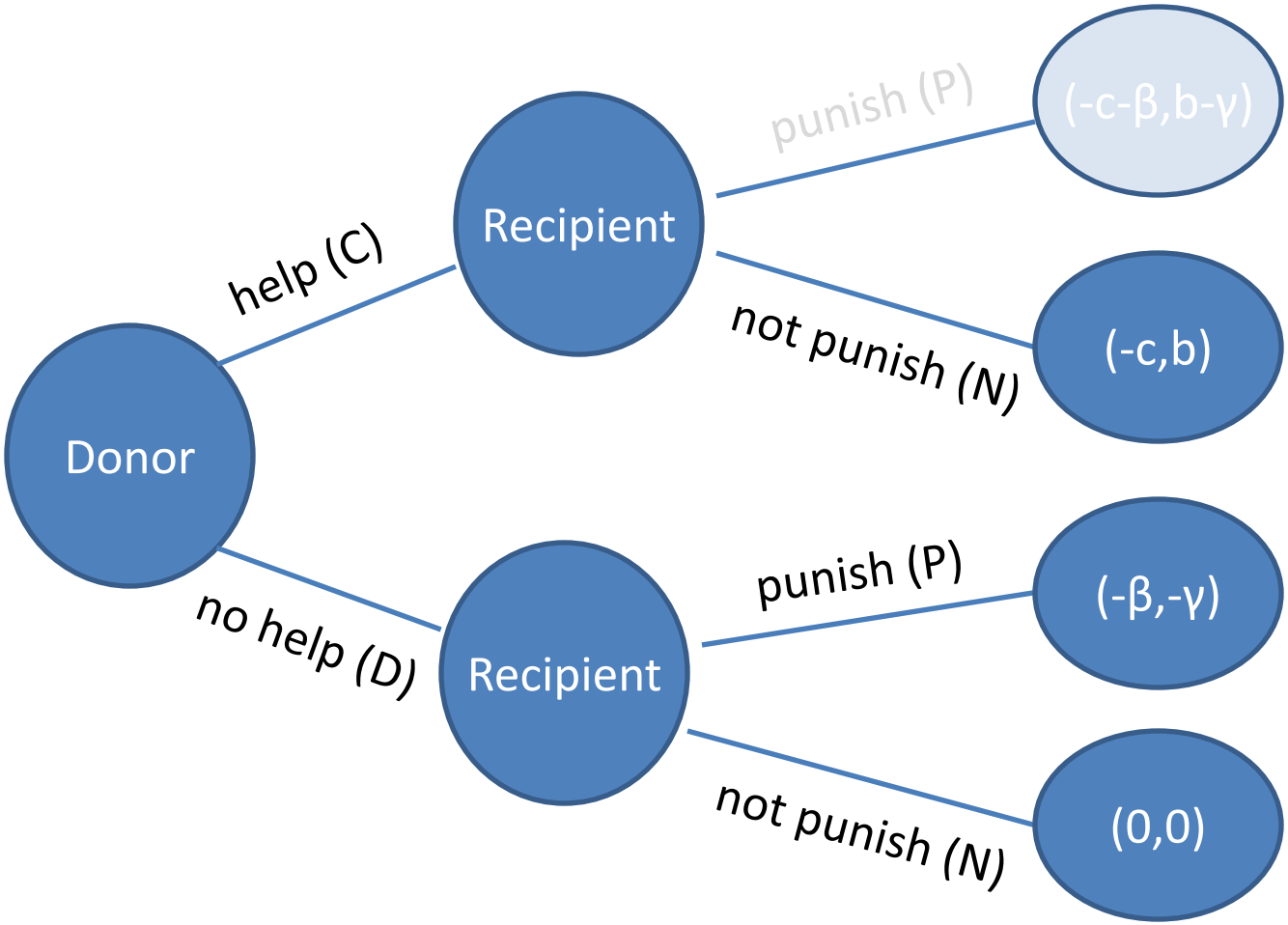


costs of helping	-c
benefit of help	b
punishment fine	-β
costs for the punisher	-γ

From the Carrot to the Stick

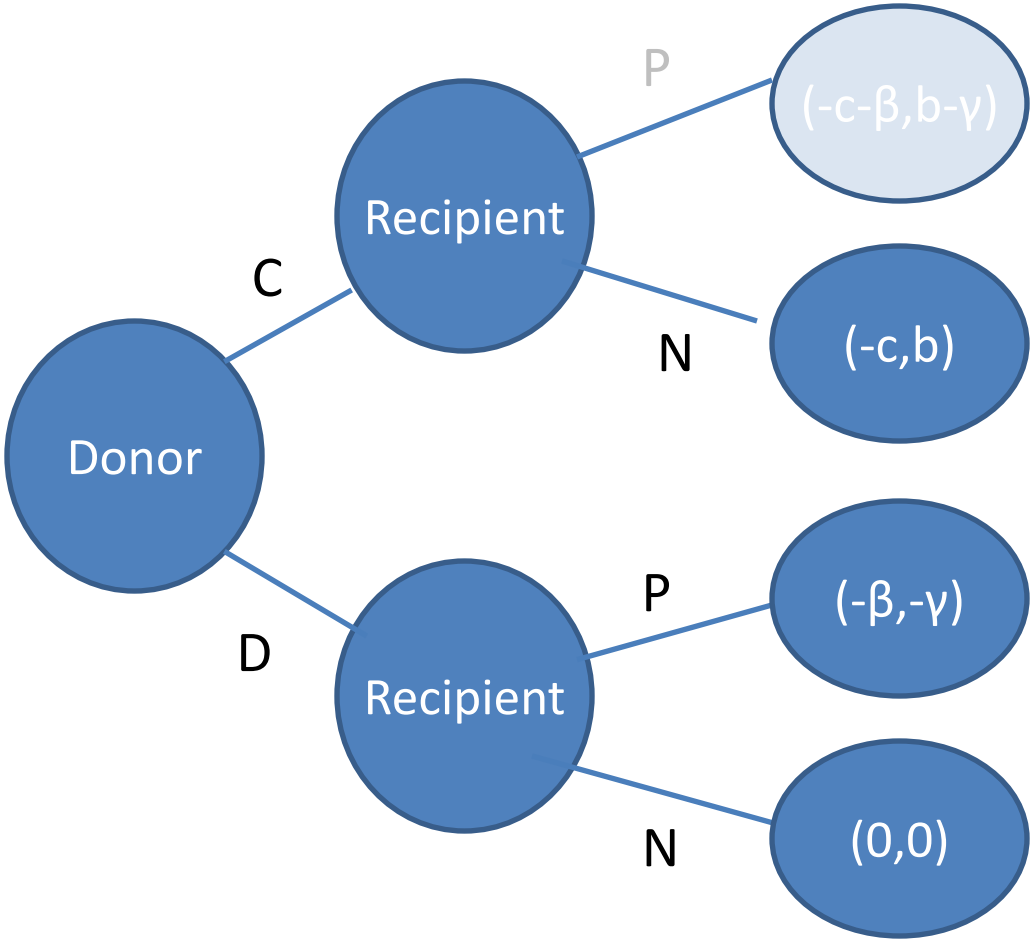
Introduction Model & Results Discussion

- Helping Game with Punishment



costs of helping	-c
benefit of help	b
punishment fine	-β
costs for the punisher	-γ

From the Carrot to the Stick



Each player has:
-2 possible strategies for the helping game (C resp. D)
-2 strategies for the incentive game (P resp. N)

In total 4 strategies:
[C,N]
[C,P]
[D,N]
[D,P]

From the Carrot to the Stick

Payoffmatrix

	$[C,N]$	$[C,P]$	$[D,N]$	$[D,P]$
$[C,N]$	$b-c$	$b-c$	$-c$	$-c$
$[C,P]$	$b-c$	$b-c$	$-c-\gamma$	$-c-\gamma$
$[D,N]$	b	$b-\beta$	0	$-\beta$
$[D,P]$	b	$b-\beta$	$-\gamma$	$-\beta-\gamma$

costs of helping	-c
benefit of help	b
punishment fine	-β
costs for the punisher	-γ

From the Carrot to the Stick

Replicator Dynamics

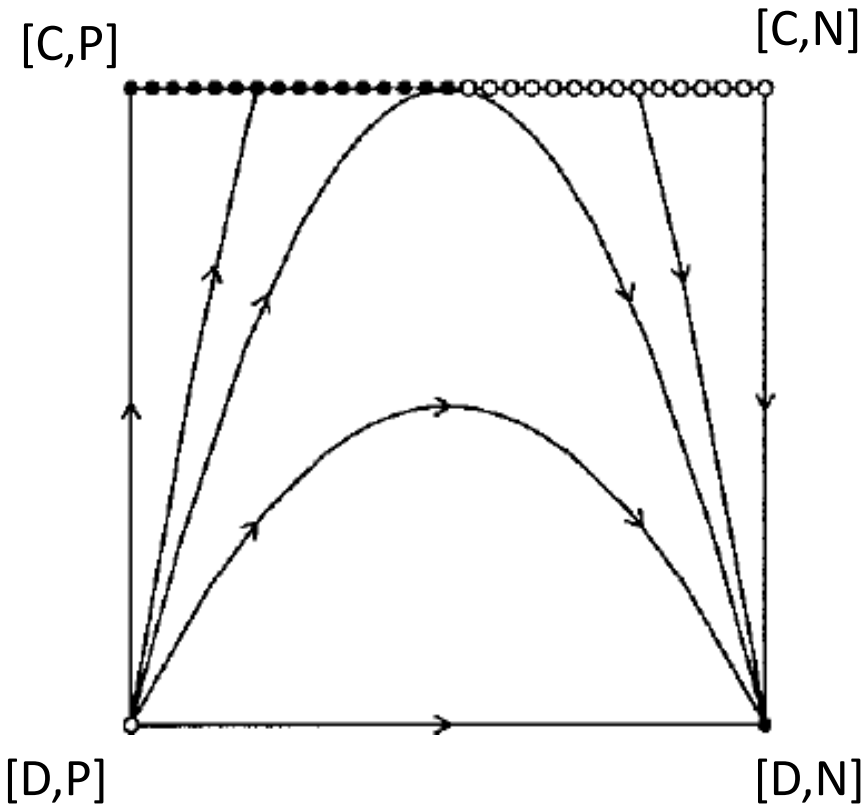
G_1=[C,P], G_2=[D,P], G_3=[D,N], G_4=[C,N]
x_i … fraction of players using G_i

Assume that x_i grows proportionally to its success in the game:

$$\dot{x}_i = x_i \cdot [(Ax)_i - x \cdot Ax]$$

From the Carrot to the Stick

Result: helping game with punishment



Sigmund et al., PNAS, 2001

- Punishment can fix cooperation, but
- Punishment itself is highly unstable

From the Carrot to the Stick

Incentives for punishment:
What can prevent the breakdown of punishment?

- Everybody who does not punish selfish individuals is punished as well (2nd order punishment) – not found in experiments

- Players are more likely to cooperate against strict co-players
  → opportunism

From the Carrot to the Stick

2 kinds of opportunism:

- An $\mathbf{O_C}$ player usually cooperates. If he knows by chance that he can get away with defection (i.e. if the co-player plays N), then he defects.
- An $\mathbf{O_D}$ player usually defects. But if he knows that he would be punished for defection, he cooperates.

μ … probability that a player knows the type of its co-player

From the Carrot to the Stick

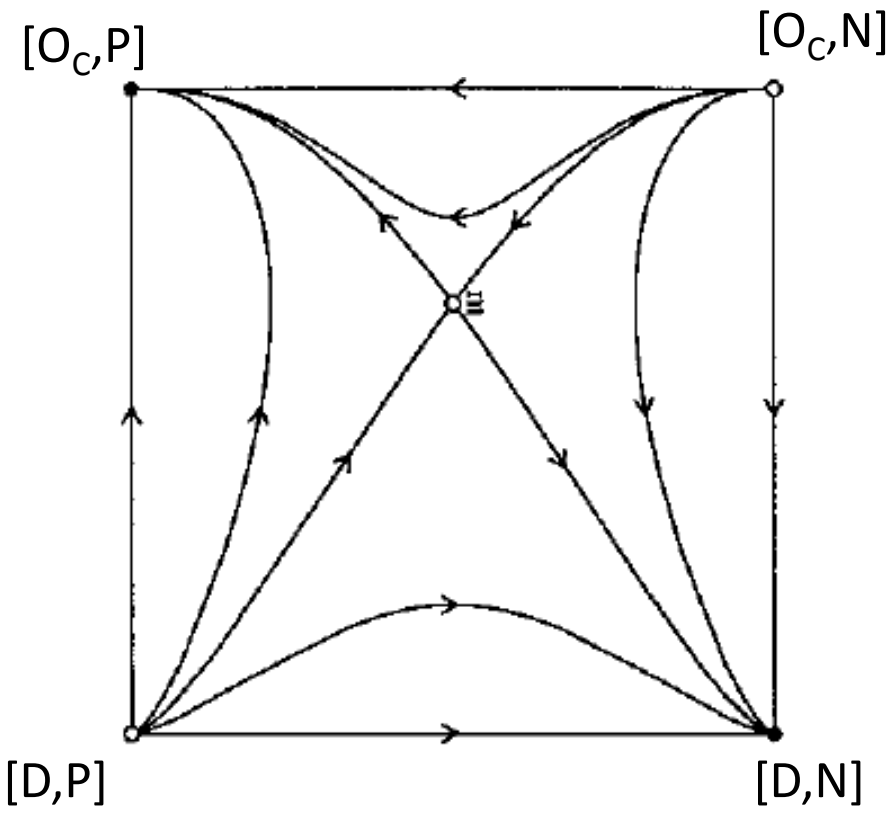
Payoffmatrix

Assumptions: c<b, c<β. $\bar{\mu} = 1 - \mu$

	$[O_C, N]$	$[O_C, P]$	$[O_D, N]$	$[O_D, P]$
$[O_C, N]$	$\bar{\mu}(b-c)$	$\bar{\mu}b - c$	$-\bar{\mu}c$	$-c$
$[O_C, P]$	$b - \bar{\mu}c$	$b - c$	$-\bar{\mu}c - \bar{\mu}\gamma + \mu b$	$-c - \bar{\mu}\gamma + \mu b$
$[O_D, N]$	$\bar{\mu}b$	$-\mu c - \bar{\mu}\beta + \bar{\mu}b$	0	$-\bar{\mu}\beta - \mu c$
$[O_D, P]$	b	$b - \bar{\mu}\beta - \mu c$	$-\bar{\mu}\gamma + \mu b$	$\bar{\mu}(-\beta - \gamma) + \mu(b-c)$

From the Carrot to the Stick

Result: helping game with punishment



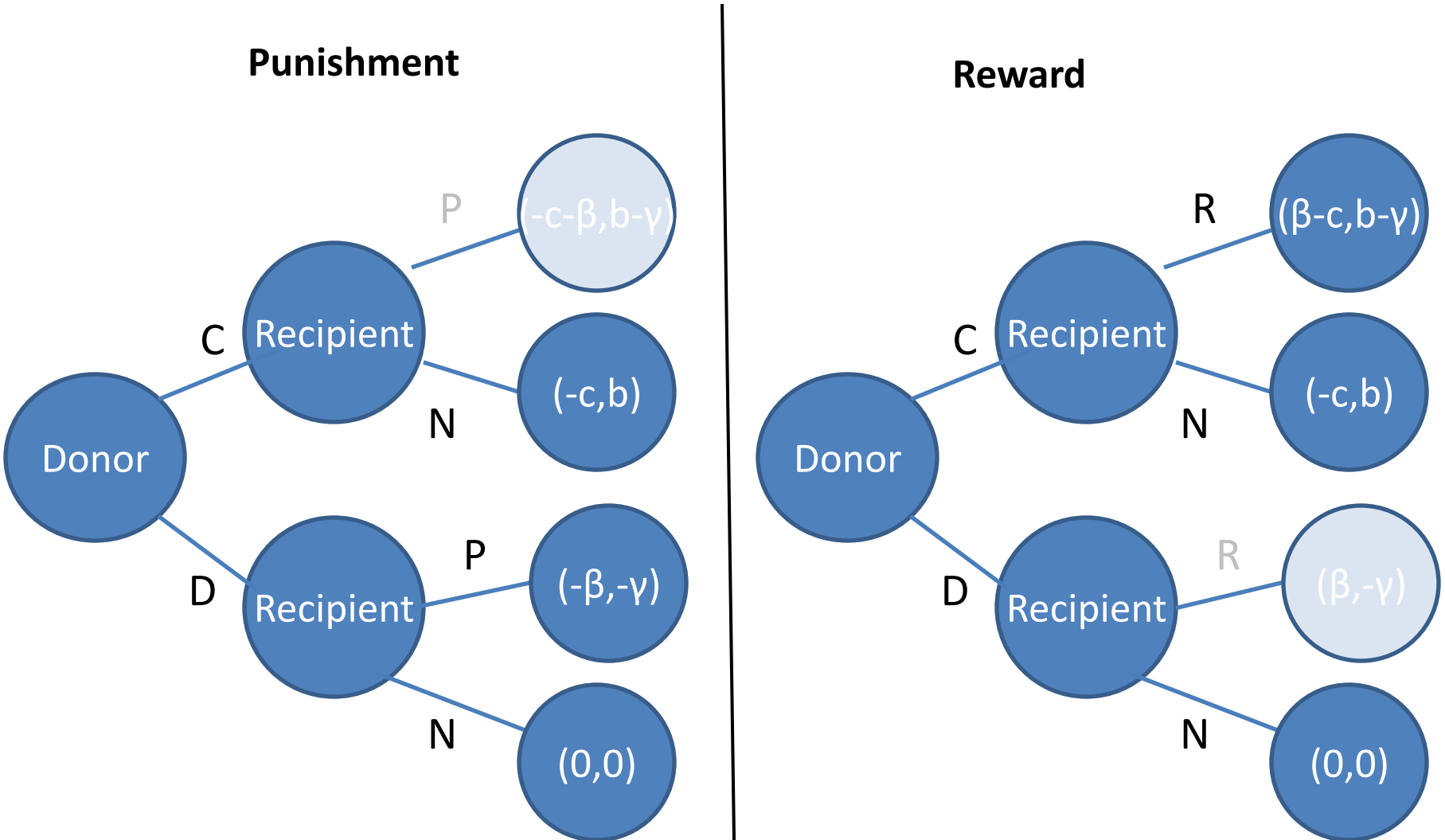
Sigmund et al., PNAS, 2001

- Bistable dynamics
- Cooperation can be fixed, but does not have a chance if initially rare

From the Carrot to the Stick

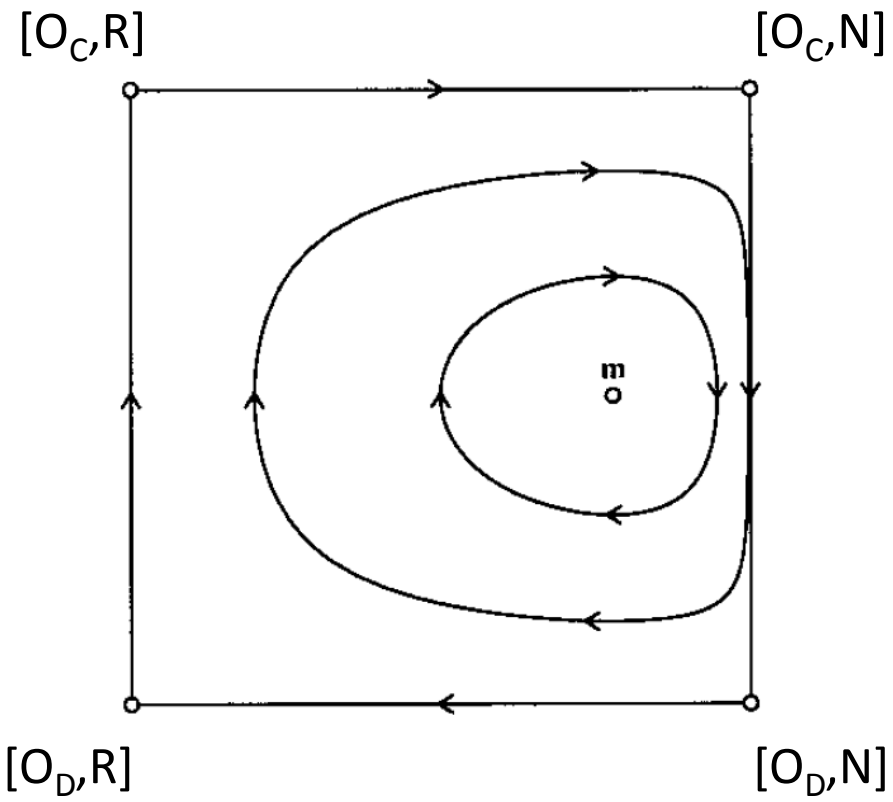
Introduction Model & Results Discussion

- Reward vs. Punishment



From the Carrot to the Stick

Result: helping game with rewards



- Cyclic dynamics
- Neither cooperation nor defection is stable

Sigmund et al., PNAS, 2001

From the Carrot to the Stick

Open questions:

What happens if both types of incentives, reward and punishment are available?

Is it possible to fix cooperation even if initially rare?

What is the more efficient / effective type of incentive?

From the Carrot to the Stick

Basic idea of our model:

- We directly follow these approaches

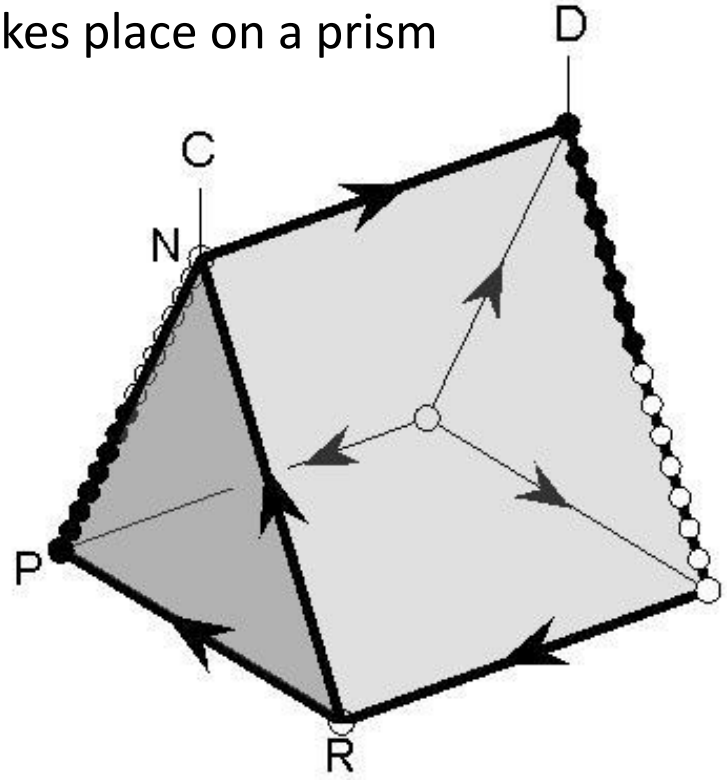
- Instead of binary options for the incentive stage (P,N or R,N), we allow all 3 strategies (P, R, N)

- For the first stage we allow 2-4 strategies (All C, All D; O_C and O_D)

From the Carrot to the Stick

2x3 role games on a prism:

Because of these invariants, the problem has essentially 3 dimensions.
→ dynamics takes place on a prism

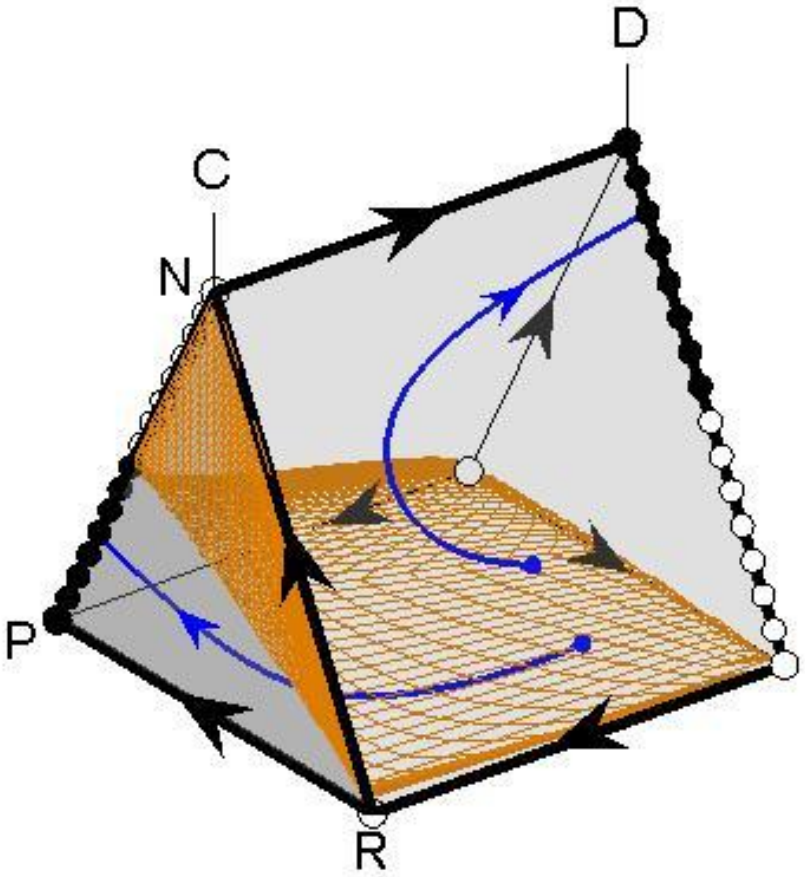


From the Carrot to the Stick

Introduction **Model & Results** Discussion

- Helping game with incentives –no information

The case of no information (μ=0)

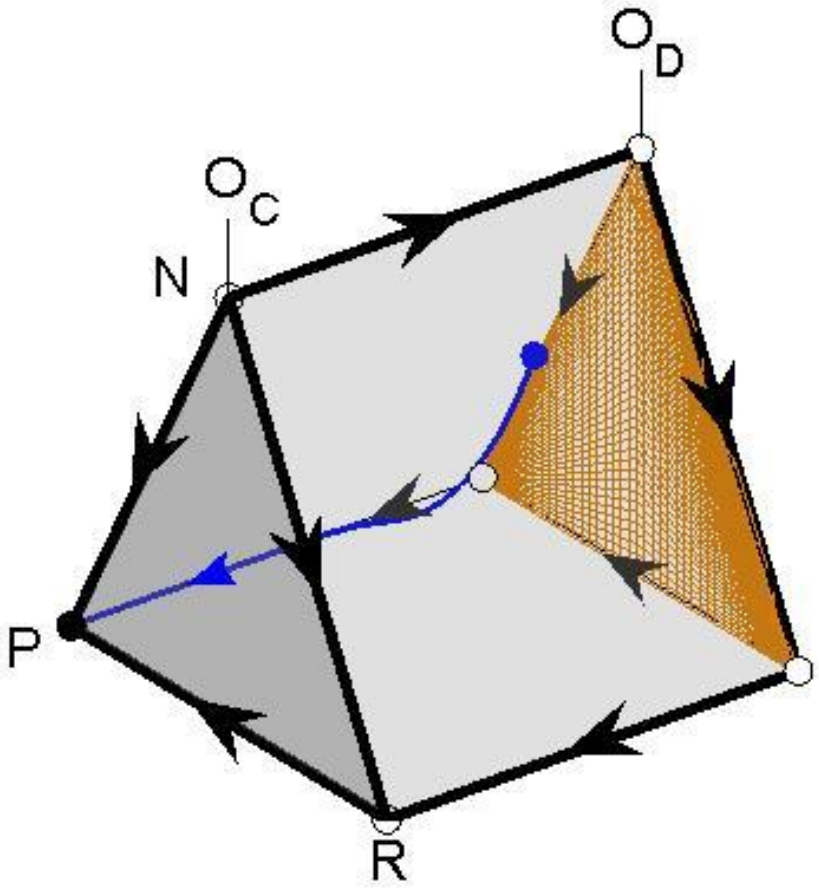


From the Carrot to the Stick

Introduction **Model & Results** Discussion

- Helping game with incentives –high information

The case of high information (μ=0.75)

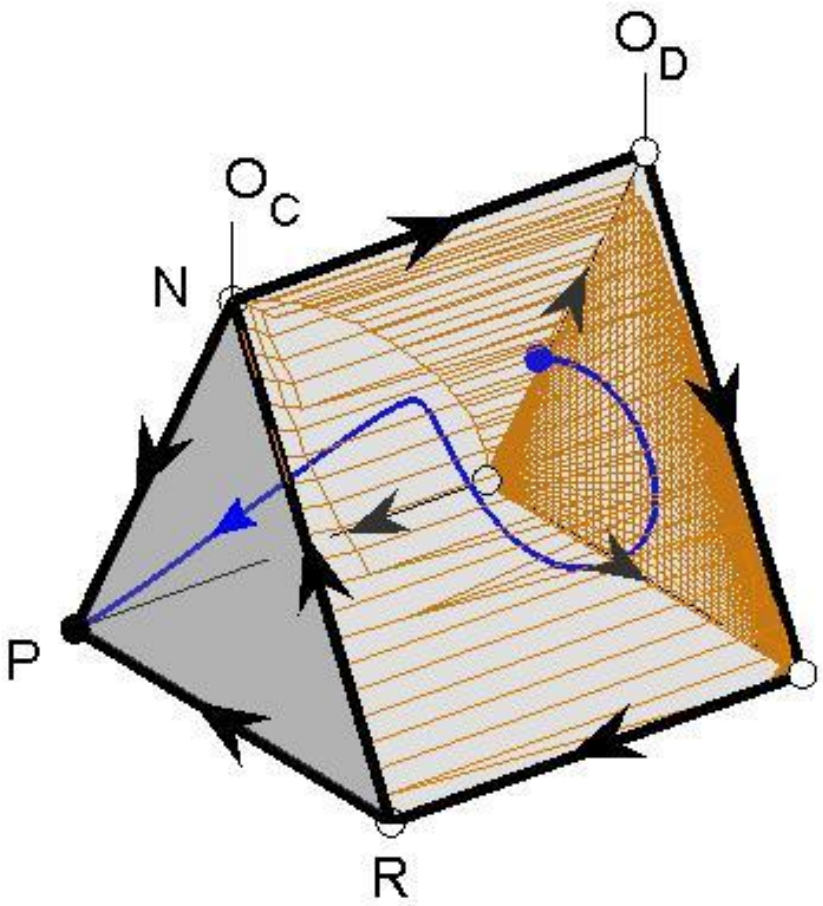


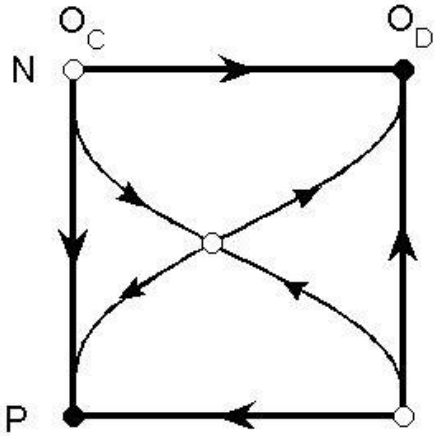
From the Carrot to the Stick

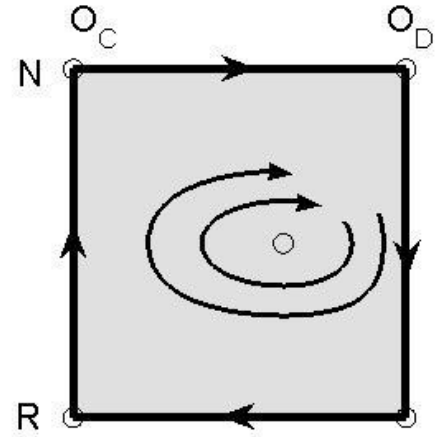
Introduction **Model & Results** Discussion

- Helping game with incentives –low information

The case of low information (μ=0.25)

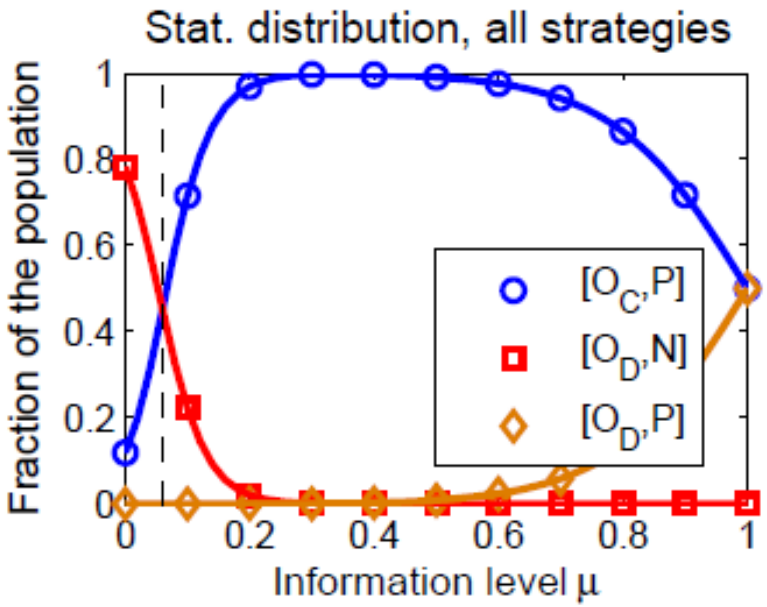


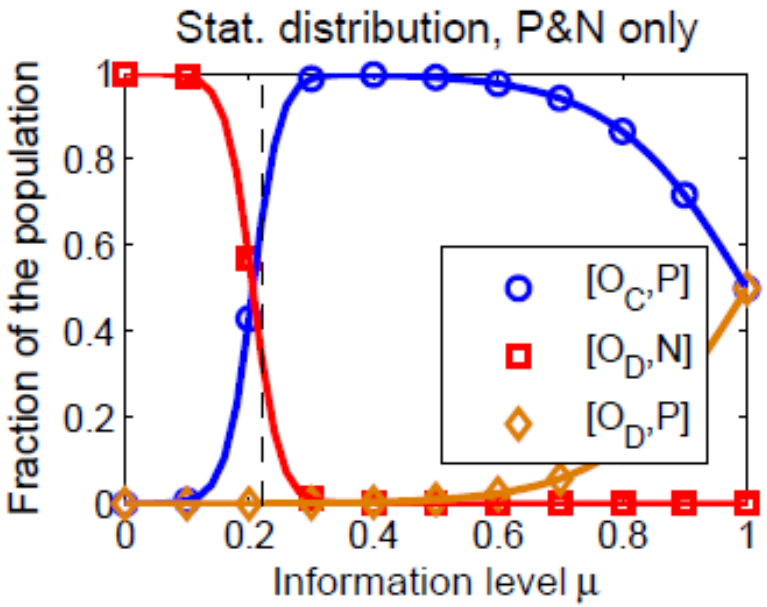




From the Carrot to the Stick

Stationary distribution (in dependence on μ)





From the Carrot to the Stick

Successful invaders into [O_D,N]

	$[O_C, N]$	$[O_C, R]$	$[O_C, P]$	$[O_D, N]$	$[O_D, R]$	$[O_D, P]$
$\mu = 0\%$	0.000	0.001	0.000	–	**0.999**	0.000
$\mu = 10\%$	0.000	0.003	0.001	–	**0.996**	0.000
$\mu = 20\%$	0.000	0.017	0.011	–	**0.970**	0.002
$\mu = 30\%$	0.000	0.068	0.040	–	**0.858**	0.034
$\mu = 40\%$	0.000	0.191	0.191	–	**0.381**	0.238

From the Carrot to the Stick

General conclusions for our model:

•Opportunism allows the evolution of cooperation

•In our model, punishment is more effective than rewards because it becomes cheaper as soon as cooperation is established

•Rewards can act as a catalyzer if the population consists of a majority of defectors

From the Carrot to the Stick

Experimental Results I: Rockenbach & Milinski, Nature 444, 2006

•Individuals played a public good game with 20 rounds

•Players had the choice to play in a game with rewards only or to chose a group with allowed punishment

•Subjects preferred the group with the opportunity of costly punishment in the second half (periods 11-20)

•Contributions are highest with both, punishment and rewards

From the Carrot to the Stick

Experimental Results II: Rand et al., Science 325, 2009

•Public good game with 50 rounds, always with the same 4 partners

•Four treatments: Control, PN, RN, RNP

•RNP leads to the highest contributions, but those who reward get the highest payoff

•Their conclusion: „Winners don't punish“ (Dreber et al., Nature 452, 2008)

From the Carrot to the Stick

Our model vs. experiments:

•Altruistic punishment:
In experiments, subjects also punish unfair behavior even if this yields no personal monetary advantage

•Incentives for providing incentives:
In our model, players could only use incentives to enforce cooperation in the helping game. In experiments, incentives may be used to control the co-player's use of incentives, e.g. counter-punishment

From the Carrot to the Stick

Thank you for your attention!

Hilbe & Sigmund, Proc. R. Soc. B, 277:2427-2433, 2010.

Questions: christian.hilbe@univie.ac.at