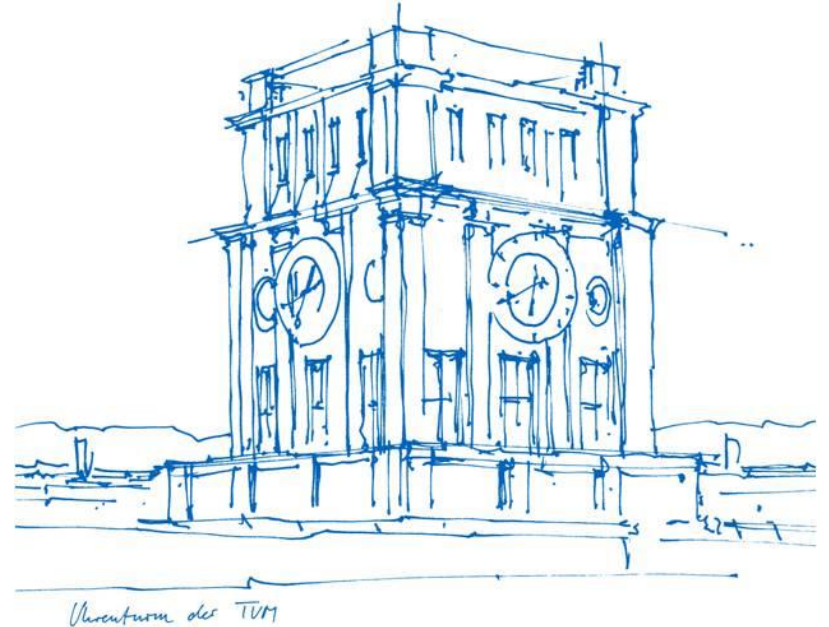


# Behavioral Economics

Prof. Dr. Sebastian J. Goerg  
Dr. Orestis Kopsacheilis

Technical University of Munich  
TUMCS for Biotechnology and Sustainability  
TUM School of Management  
Department of Economics and Policy

Winter 2020/21



## Course Overview

- I. What is Behavioural Economics
- II. Principles of Experimental Economics
- III. The Standard Economic Model: Consumer Theory
- IV. Reference dependence & departures from the standard model
- V. Decisions Under Risk and Uncertainty(I & II)
- VI. Intertemporal Choice
- VII. Interaction with others: Game Theory
- VIII. Interaction with others: Beh. Game Theory & Social Pref/ces

# Today

## VIII. Interaction with others: Beh. Game Theory & Social Pref/ces

### A. Limited strategic thinking

- P-beauty contest, Level-k reasoning, Cognitive Hierarchy

### B. Multiple equilibria and coordination

- Focal points, Schelling's salience, Pareto & risk dominance

### C. Social preferences

- Social dilemmas, Conditional Cooperation, Intentions

# Keynesian beauty contest

Consider a beauty contest where judges are rewarded for selecting the most popular faces among all judges, rather than those they may personally find the most attractive.

*"It is not a case of choosing those [faces] that, to the best of one's judgment, are really the prettiest, nor even those that average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees."*

- Keynes, General Theory of Employment, Interest and Money, 1936



John Maynard Keynes (1883-1946)

# P-beauty contest, application: stock market

- You own stock from a company and its price is currently trending upwards.
- You know that its current price is above its fundamental value so it's bound to deflate soon (bubble).
- Bubbles typically burst when stock owners start selling massively.
  
- You want to sell before it crushes BUT you also don't want to sell while the price is climbing.
- Assume that:
  - You need to make a decision re. when to sell the stock between today and up to 100 days from now.
  - Let  $D$ , be the average number of days that other investors choose to delay selling and that the stock keeps on appreciating  $\frac{2}{3} * D$ .
  - You want to sell as close to  $\frac{2}{3}$  times the average chosen by others as then you manage to cash in at the max. value of the stock's price.
- How many days must you wait before you sell?

# P-beauty contest, application: stock market

- In keeping with Keynes's analogy, it has been suggested that this kind of game captures the dynamics of real markets and can explain bubbles in stock and real-estate markets.
- Even if all investors know that the market will ultimately crash, and that the unique Nash equilibrium strategy is to exit the market, they might assume that others will continue to buy for just a little while longer.
- As long as individual investors think that they can stay one step ahead of the competition and exit the market just before everybody else does, they will want to continue to buy. In the process, of course, they will drive prices even higher.

# P-beauty contest

You are playing this game with your classmates from the Behavioral Economics course. Every student guesses a number between 0 and 100. The winner is the person who will guess the number closest to  $2/3$  the average of guesses. Please, make a guess by sliding the bar below accordingly.

0      10      20      30      40      50      60      70      80      90      100

Make a guess



# NE: Iterated elimination of dominated strategies

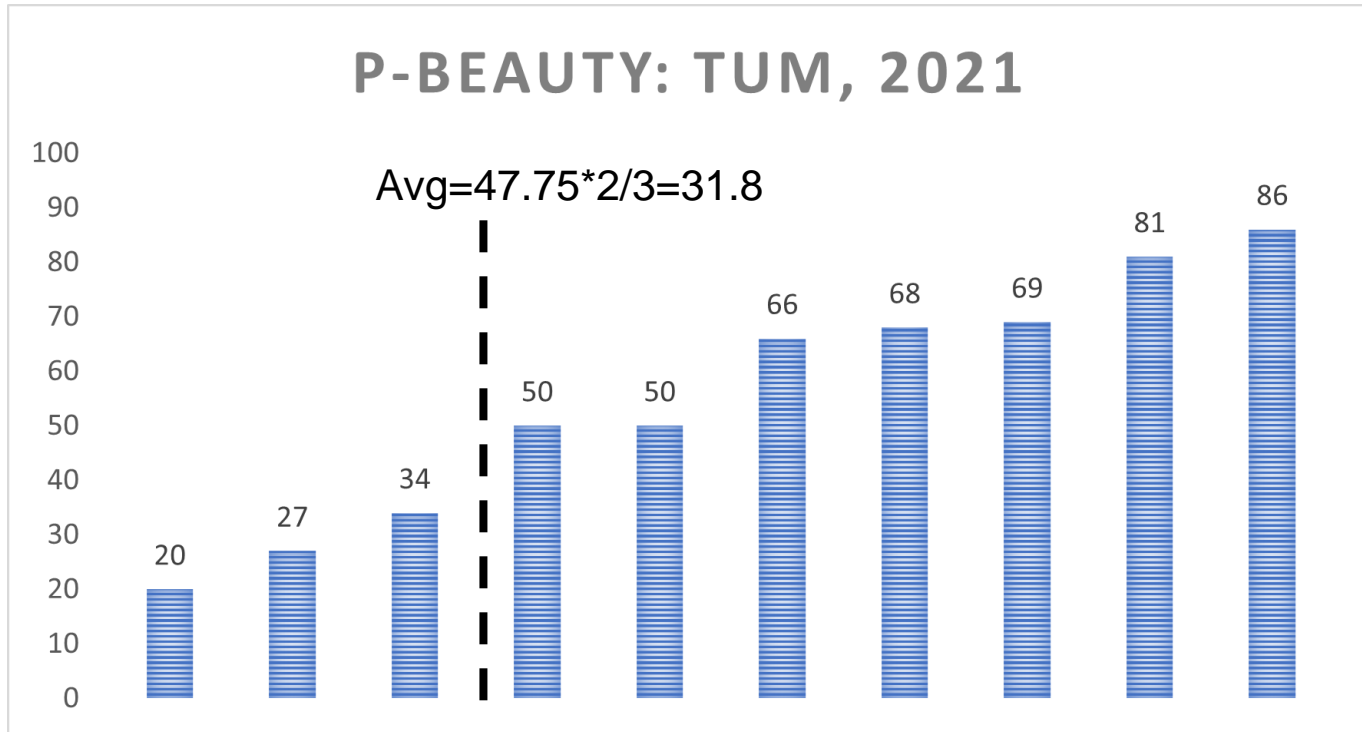
- First, notice that anyone who plays this game strategically (instead of just picking a random number) will likely avoid numbers higher than  $100 * \frac{2}{3}$ 
  - Therefore, any number higher than 66 is strictly dominated by 66.
  - Of course, you may reason that it is unlikely that everyone else waits 100 days. The same principle applies even if you make a different step-1 guess.
  - This is often referred to as **Step 0** of reasoning
- We have assumed common knowledge (everyone knows this and everyone knows that everyone knows this...), so everyone expects that everyone will pick a guess  $\leq 66$ .
- **Step 1** of reasoning: Since no one is expected to wait for more than 66 days, all guess higher than  $\frac{2}{3} * 66 = 44$  are strictly dominated by 44...
  - We say that a strategy to wait more than 44 days is one step dominated by a strategy to wait 44 days.
- But the reasoning doesn't stop there. If you believe that no one will ever play a one-step dominated strategy then...



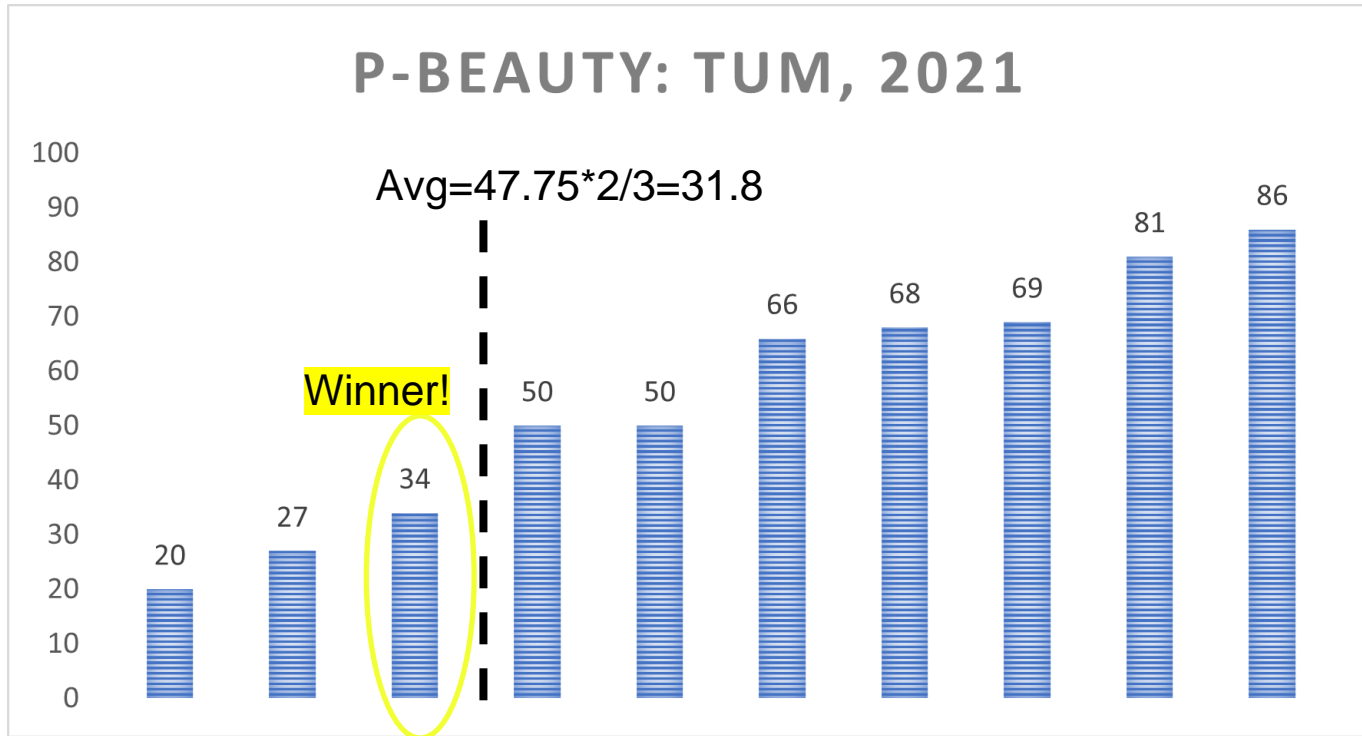
# NE: Iterated elimination of dominated strategies

- **Step 2:** 44 is now the new upper boundary, so guesses higher than  $\frac{2}{3} * 44 = 29$  are two-step dominated by a strategy to wait 29 days
- **Step 3:** ...  $\frac{2}{3} * 29 = 19.3$
- .
- .
- .
- NE= wait 0 days. Therefore, according to formal theory, only one strategy survives iterated deletion of dominated strategies: 0 (i.e. selling today).
- Is this what people do?

# Results from p-beauty ( $p=2/3$ ) contest, TUM, 2021

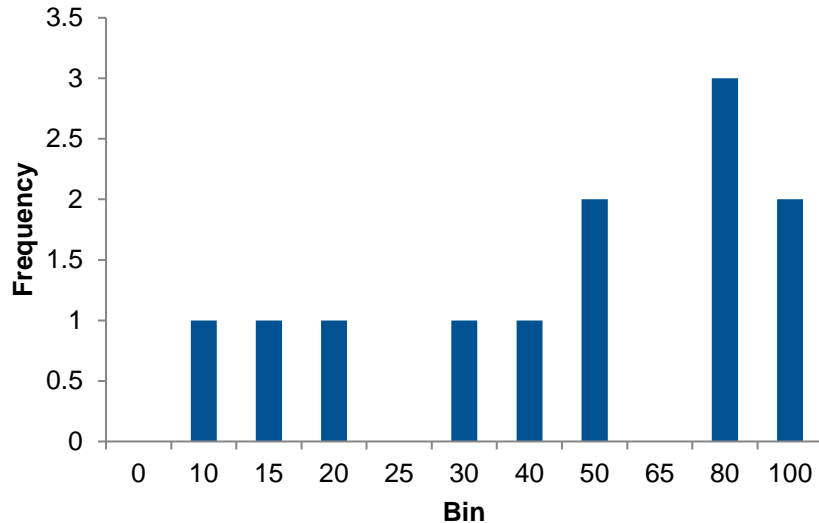


# Results from p-beauty ( $p=2/3$ ) contest, TUM, 2021



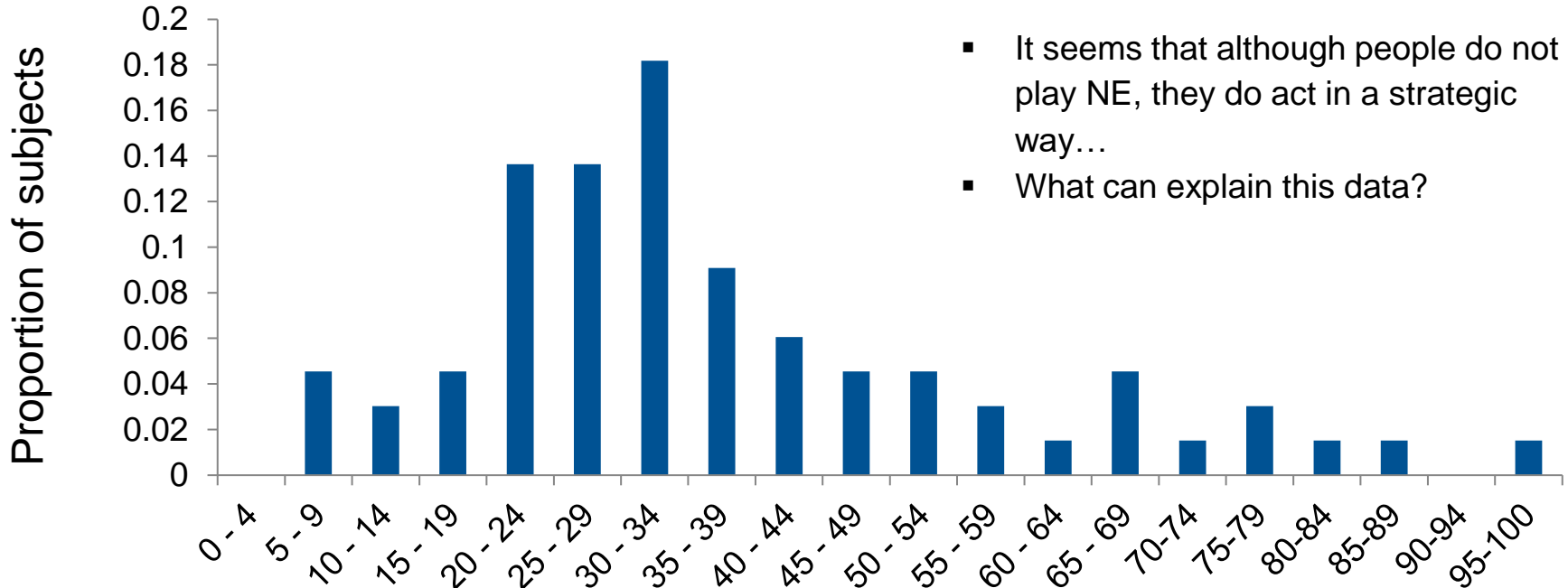
# Results from p-beauty ( $p=2/3$ ) contest, TUM, 2021

## Histogram



1. Some people said more than 66 (a strictly dominated strategy)
2. No-one said 0.
  - Does this mean that people were wrong?
  - In order to win, it's not enough that you understand how to "solve" the game for its NE. It is also important to understand what others think.

# P-beauty-contest ( $p=2/3$ ) responses in Nagel, 1995



- It seems that although people do not play NE, they do act in a strategic way...
- What can explain this data?

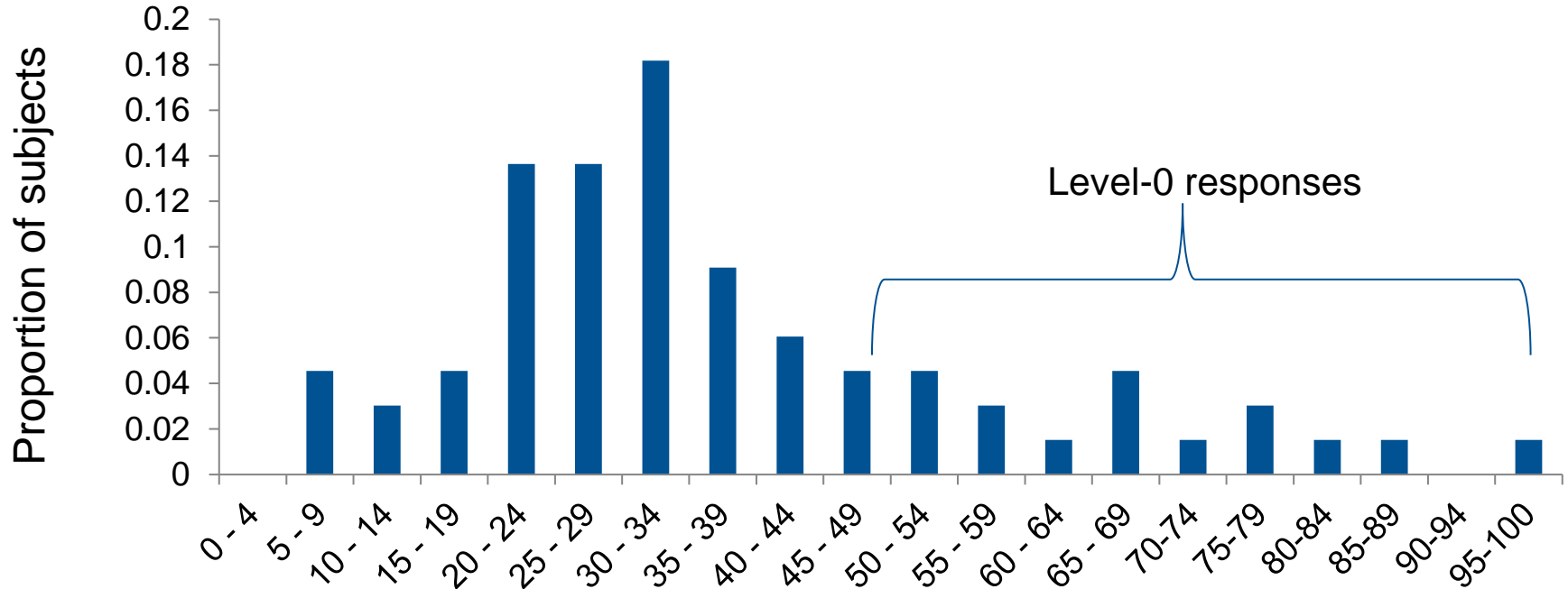
# Level-k thinking

- A theory of limited strategic thinking (Stahl and Wilson, 1995; Bacharach and Stahl, 2000)
- Basic idea: someone forms an expectation of what others will do and tries to reason 'one step ahead of them'.
- Level-0 type: chooses a number randomly
- Level-1 type:
  - Thinks that everyone is Level-0, so the average number must be 50.
  - Takes the  $2/3$  of that average: 33
- Level-2 type:
  - Thinks that everyone else is Level-1
  - Takes the  $2/3$  of  $33 = 22$
- Level-3 type:  $2/3$  of  $22 \approx 14$
- ...

# Level-k thinking

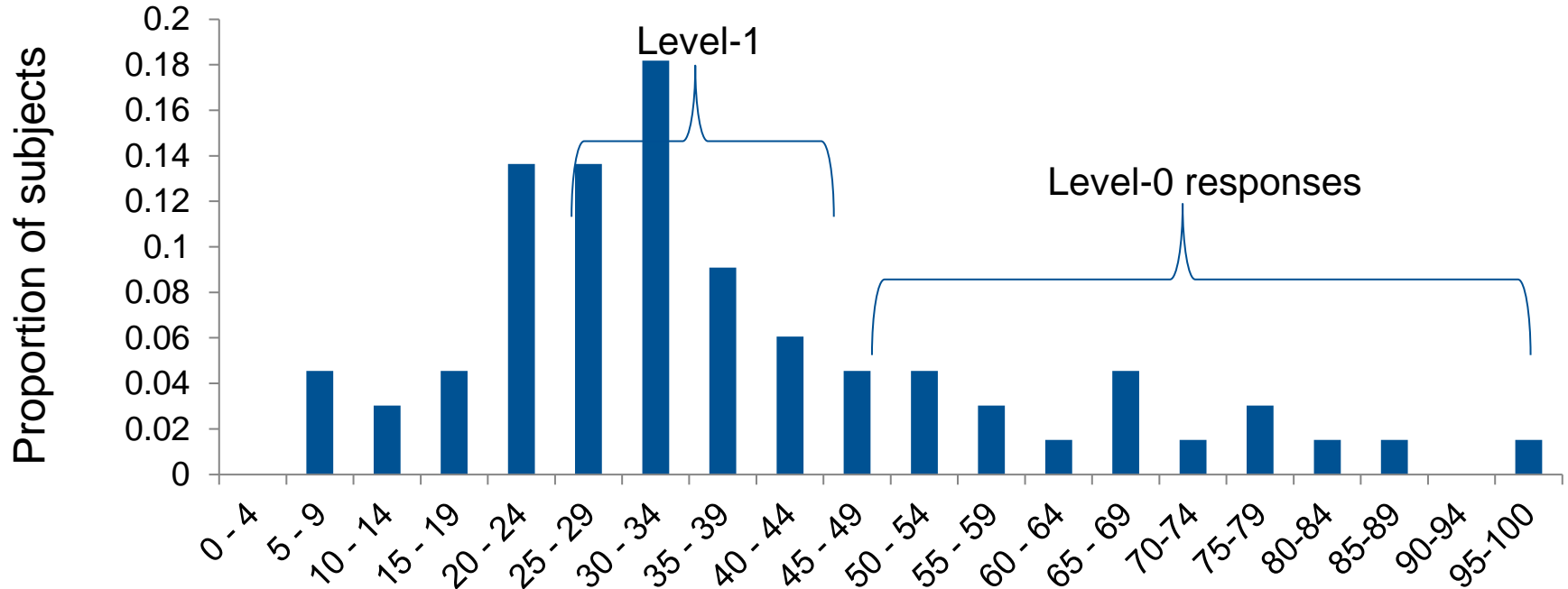
- Level-k: thinks that everyone is reasoning at k-1 level.
- The Nash equilibrium of this game, where all players choose the number 0, is thus associated with an infinite level of reasoning.
- Empirically, in a single play of the game, the typical finding is that most participants can be classified from their choice of numbers as members of the lowest Level types 0, 1, 2 or 3, in line with Keynes' observation.

# Beauty-contest ( $p=2/3$ ) responses in Nagel, 1995

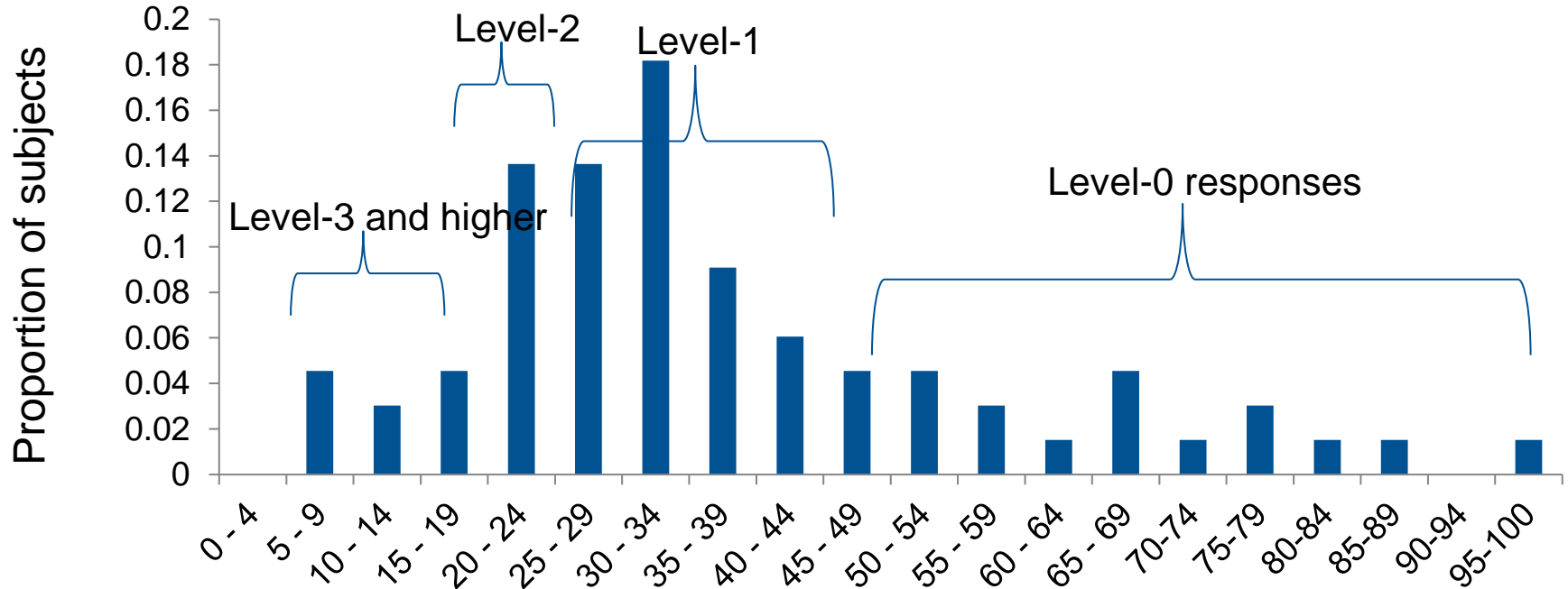




# Beauty-contest ( $p=2/3$ ) responses in Nagel, 1995



# Beauty-contest ( $p=2/3$ ) responses in Nagel, 1995



# Level-k, Criticism 1: What do Level-0s do?

- What do Level-0 players do?
- In our example earlier we assumed that Level-0 picks a number randomly.
- This is a crucial assumption as Level-1 types base their expectations on that and by extension, all predicted types after that...
- But other possibilities are also possible
- It has been suggested that Level-0 pick focal strategies (in this case numbers).
- Focal point: something that stands out... In this case, round numbers like 50 or 100...
- Another suggestion is that Level-0 play the NE...
- Level-k theory gives different predictions depending on what we assume as Level-0 behavior.
- That's a problem...

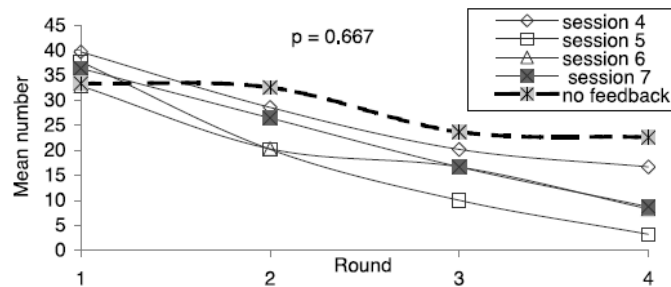
# Self-reports vs. Observed behaviour

- This is a comment on the methodology of experimental economics.
- Remember, in the last slide we asked “what motivates Level-0 reasoners?”
- Notice, we cannot know just by observing the data. All we see is a number per person. Intentions are inferred. We cannot look inside participants brains.
- Therefore, there is no way to tell for sure if someone who said “22” picked a random number (Level-0) or exercised a level-2, k-reasoning. Moreover, we cannot tell if someone who picked “50” did so because that number was standing out or because he picked a number at random.
- Asking is one option but economists are sceptical about it for several reasons.
- One reason is that these strategies may be sub-conscious and people cannot describe them.
- Another reason is that self-reports cannot be monetarily incentivised in incentive-compatible way. So, people may “lie”. For example, if we ask how altruistic they are, people might say “very much”. But, will they contribute to a charitable cause if endowed with money, or will they keep all to themselves?
- With technology advancements, we can make some progress in this direction (e.g. mouse-lab, eye-tracking technology, f-MRI scans...).

# Level-k, Criticism 2: lack of sophisticated beliefs

- Remember, a level-k type thinks that everyone else is level-(k-1) thinkers.
- Beliefs become increasingly implausible as k gets higher...
- For example, a Level-3 reasoner would expect that the average in the p-beauty contest ( $p=2/3$ ) would be close to 10... This is very low...
- Perhaps, a more plausible idea, is that a Level-k reasoner would think that there is a distribution of types. Some are Level-2, some are Level-1 and some are Level-0.
- This is the idea behind the Cognitive Hierarchy Model (Camerer, Ho and Chong, 2004)

# Repetition in the p-beauty contest



The mean number chosen when playing a p-beauty contest four consecutive times, for each of the experimental sessions.

Sources: Nagel (1995) and Weber (2003).

- Sessions 4-7: participants were given feedback about the winning guess after each round.
- Most people guessed too high initially and adjusted downwards in ensuing rounds, approaching in some cases the NE.
- Interestingly, there was some learning even in the “no-feedback” condition, suggesting that deliberation alone can provide learning effects.

# Today

## VIII. Interaction with others: Beh. Game Theory & Social Pref/ces

### A. Limited strategic thinking

- P-beauty contest, Level-k reasoning, Cognitive Hierarchy

### B. Multiple equilibria and coordination

- Focal points, Schelling's salience, Pareto & risk dominance

### C. Social preferences

- Social dilemmas, Conditional Cooperation, Intentions

# Strategic thinking in pure coordination games

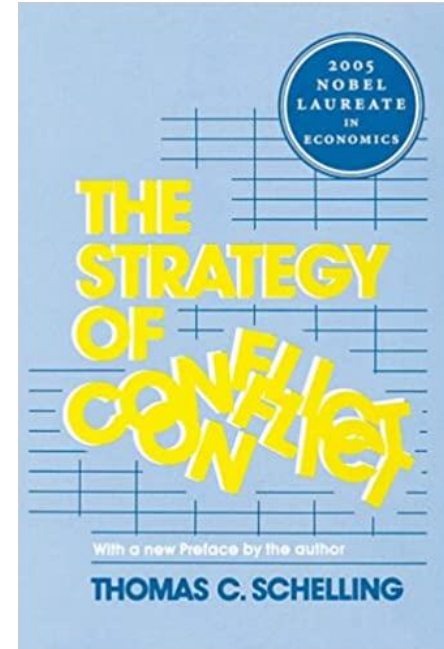
- Consider again the p-beauty contest, but this time, the winners are those who choose the most popular answer.
- This is a pure coordination game. Remember from last lecture the café-coordination game and the “which tire” problem.
- According to standard theory, the NE are as many as the strategies in the game:  $\{0, 1, \dots, 100\}$
- Without the ability to communicate, it might seem pretty hopeless to coordinate.
- Should one choose a strategy completely at random or are there ideas that might improve chances of win (coordination)?



# Focal points and Schelling's salience

- Common finding: players of one-shot coordination games are way more successful than classical game theory predicts. Why?
- Thomas Schelling (*The Strategy of Conflict*, 1960) asked people to imagine that they had to meet someone in New York, but are unable to communicate with that person.
- Where should they go in the hope of meeting?
- New York is a big city, so this seems like a difficult problem but most people named the same place: **Central Station**

*Central Station is a **focal** place for coordination. But what makes a place focal?*



# Salience

- **Primary salience** hypothesis: people choose the option that is most salient to them. Because people are in many ways similar (e.g. often share cultural background, etc.) people coordinate by accident. Under this hypothesis, people coordinated in Grand Central Station, because that's their favourite place in NY
- **Secondary salience** hypothesis: people expect others to use primary salience and so choose the option they think will have primary salience for others. For example, your favourite place in NY might be the Museum Of Modern Art (primary salience). But, you recognise that your taste is not the most popular and choose the Central Station instead, that you recognise as a more popular destination for most other people.
- **Schelling salience** hypothesis. The Schelling salience hypothesis is that people ignore what is primary or secondary salient and look for some key or clue to how to coordinate. For example, Grand Central Station may not be the favorite place of anyone, but somehow stands out as the best choice

# Saliency, level-k and team reasoning

- Secondary saliency is a lot like L-1 reasoning in level-k theory where Level-0 assumes that people choose according to primary saliency (no strategic thinking).
- Schelling saliency differs. There, the individual asks to himself ‘What should we do as a team?’ and acts upon the answer in the expectation that other members of the group think and behave analogously. In other words, people ignore what is primary or secondary salient and look for some key or clue to how to coordinate. This idea has been formalised using the concept of *team reasoning* (Sugden, 1993; Bacharach, 1999).

# Saliency, level-k and team reasoning

- Mehta, Starmer and Sugden (1994):
  - a. Write down any year, past, present, or future
  - b. Write down any positive number
  
- Condition 1: No incentive to coordinate
- Condition 2: Incentivised to match others responses
  
- Idea: the answers of those with no incentive to match should tell us what was primary salient.
  - If people use primary saliency to coordinate then responses between “no incentive to match” and “incentive to match” should not be very different.
  - If people use secondary saliency, then popular answers from the unincentivized condition would become focal in the “incentivised to match” condition.
  - If people use team reasoning, then the coordinating answer will not necessarily have primary saliency

*Table 6.4* The answers subjects gave, and the proportion giving each answer, to two questions. The answers depend on whether the subjects did or did not have an incentive to match with others.

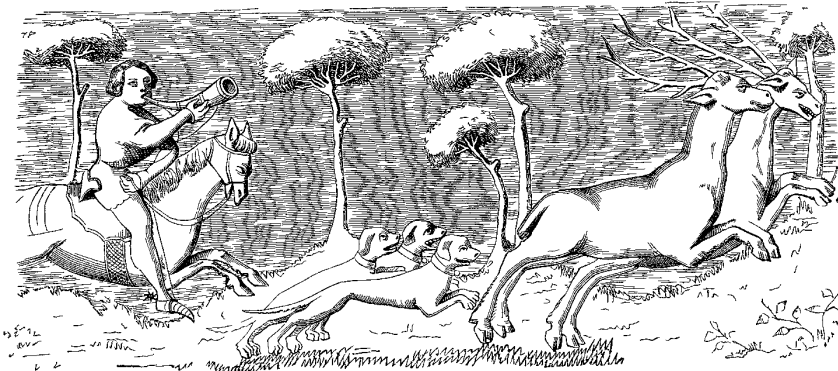
<i>Question</i>	<i>No incentive to match</i>		<i>Incentive to match</i>	
	<i>Response</i>	<i>Proportion</i>	<i>Response</i>	<i>Proportion</i>
'Write down any year'	1971	8.0	1990	61.1
	1990	6.8	2000	11.1
	2000	6.8	1969	5.6
	1968	5.7		
'Write down any positive number'	7	11.4	1	40.0
	2	10.2	7	14.4
	10	5.7	10	13.3
	1	4.5	2	11.1

*Source:* Mehta, Starmer and Sugden (1994).

- People don't use primary salience to coordinate: The variance of responses is much higher in the unincentivized condition.
- It seems that an ability to coordinate can come from either secondary salience or Schelling salience. Which fits better seems to change depending on the context.
  - Answers in the “year-question” seem to favour secondary salience: the coordinating response: “1990” was among the most popular picks in the no incentive condition but...
  - the “number-question” is suggestive of team reasoning as “1” was among the least popular responses in the “no incentive” condition

# Payoff dominance

		Player 2	
		Stag	Hare
Player 1	Stag	100,100	0,0
	Hare	0,0	80,80



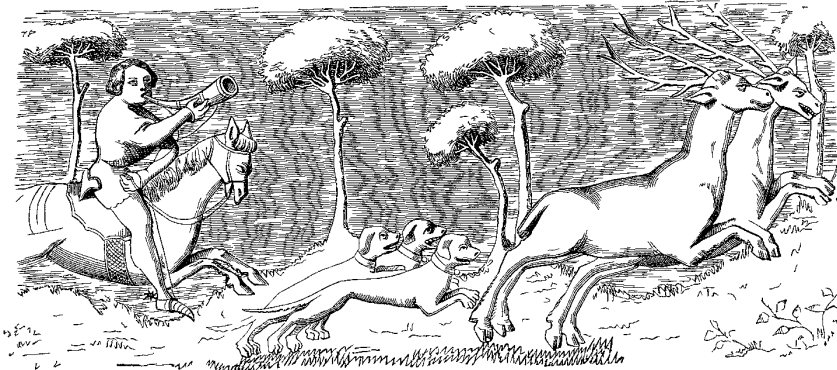
# Payoff dominance

Player 1

	Stag
Stag	
Hare	

Player 2

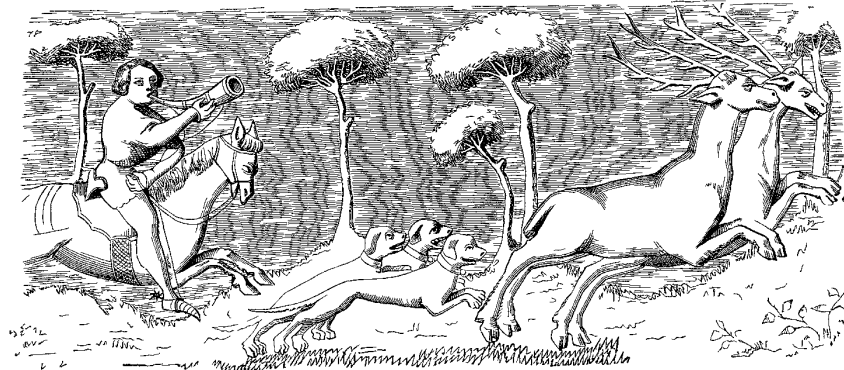
	Stag	Hare
Stag	100, 100	0, 0
Hare	80, 80	0, 0



- NE in pure strategies: (Stag, Stag) and (Hare, Hare)
- (Stag, Stag) is the payoff dominant Nash Equilibrium

# Payoff dominance vs. risk dominance

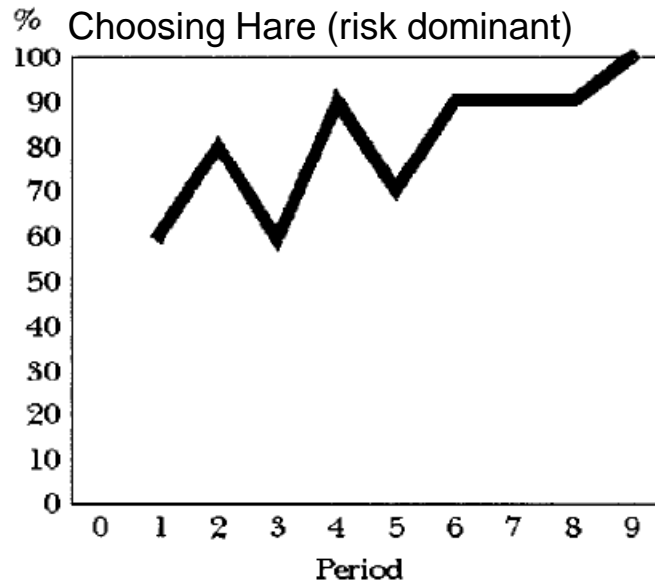
		Player 2	
		Stag	Hare
Player 1	Stag	100,100	0, <b>80</b>
	Hare	<b>80</b> ,0	80,80



- In this form (Hare, Hare) is the risk dominant equilibrium.
- Choosing “Hare” guarantees 80, irrespective of what the other player does.



# Empirical evidence for risk dominance



- Stranger's design: new match after every round.
- Initially, people try to coordinate on the payoff dominant outcome.
- But... after a couple of negative outcomes they give up.
- The risk dominant equilibrium takes over.

# Interpreting 'Stag-hunt'

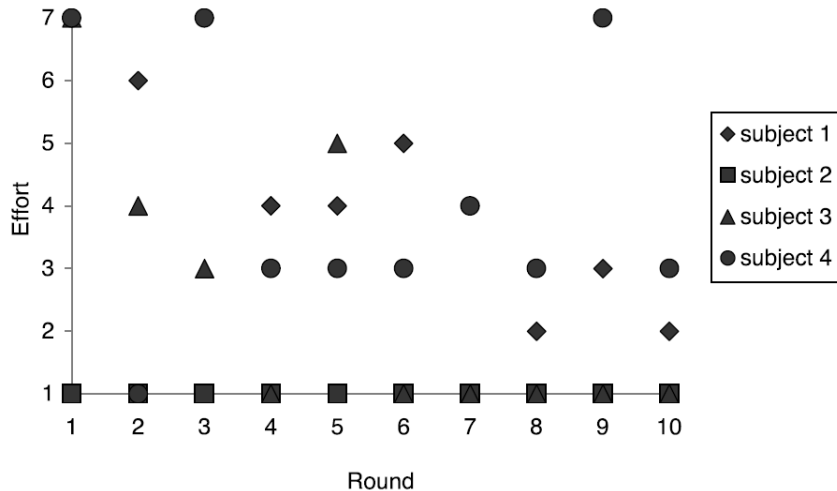
- The rationale behind it is that communal action yields a higher return if all players combine their skills, but if the other player's intentions to contribute in this team effort are unknown, then going for the payoff-inferior but safer option might be the better individual strategy.
- Message: collective action might fail in the absence of credible commitments.
- Can the result be reversed? Can people coordinate on the payoff dominant equilibrium?

# Minimum effort game

<i>Your choice of effort</i>	<i>Smallest choice of effort in the group</i>						
	7	6	5	4	3	2	1
7 (high)	13	11	9	7	5	3	1
6		12	10	8	6	4	2
5			11	9	7	5	3
4				10	8	6	4
3					9	7	5
2						8	6
1 (low)							7

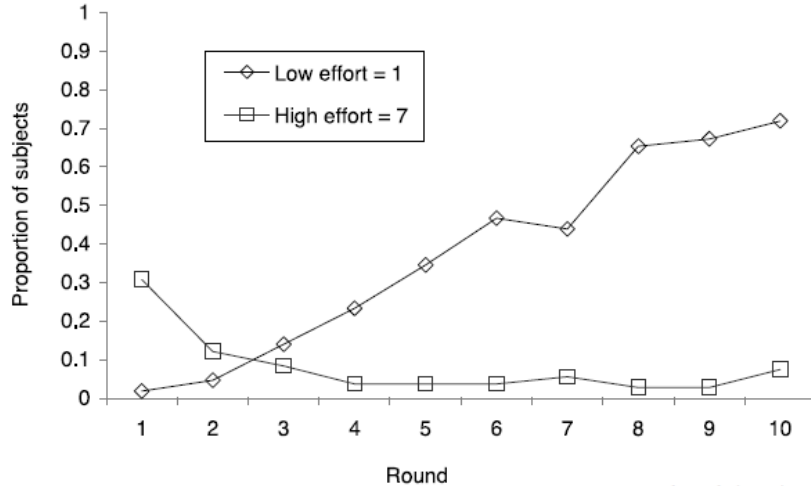
- Introduced by Van Huyck, Battalio and Beil (1990)
- Players choose effort level/ Higher effort is costly
- Minimum effort determines productiveness of team
- 7 pure-strategies NE: where everyone chooses the same effort.
- Ideally, people coordinate on high effort but...
- The safest thing for the individual, however, the risk dominant Nash equilibrium, is to put in low effort. This guarantees a payoff of 7.
- Choosing higher effort is risky because it only takes one member of the group to choose low effort for the high effort to be wasted and costly.

# Minimum effort game



- Consider this example from Gillet, Cartwright and Van Vugt (2009).
- A group of 4 is playing the same game for 10 rounds.
- The high effort of Subjects 1, 2 and 3 is going to waste as player 2 always chooses the lowest amount of effort.
- Therefore, player 2 receives 7 while everyone else less than that.

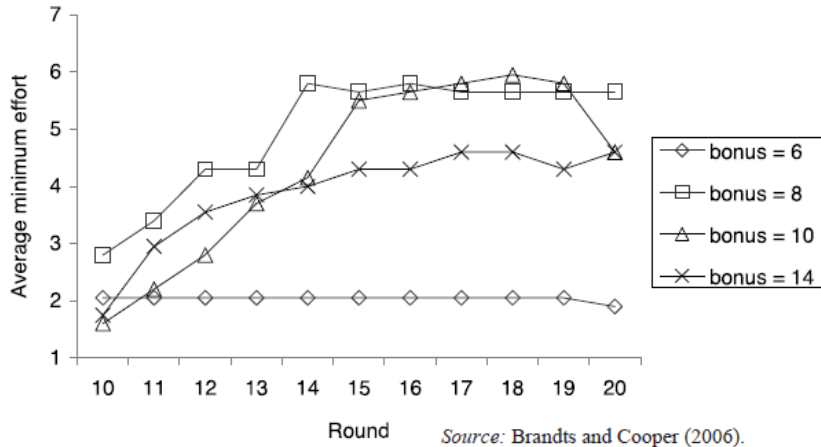
# Minimum effort game



Source: Van Huyck et al. (1990).

- The proportion of players choosing the low effort strategy is increasing.
- The proportion of players choosing the high effort is decreasing.
- **Idea:** repeat the experiment but introduce an additional incentive for coordinating in high effort
- You can think of this as a manager stepping in to boost productivity

# Minimum effort game – manager’s intervention



- The manager introduces a bonus for coordinating at the high effort strategy (7).
- Initially, the coordinating bonus was equal to 6 (coordinating at effort level=7 yields final payoff:  $7+6=12$ ). So the “bonus=6” condition is the control where no change took place.
- Notice how effort level was always low in the beginning, before the intervention.
- After the intervention effort goes up!

# Today

## VIII. Interaction with others: Beh. Game Theory & Social Pref/ces

### A. Limited strategic thinking

- P-beauty contest, Level-k reasoning, Cognitive Hierarchy

### B. Multiple equilibria and coordination

- Focal points, Schelling's salience, Pareto & risk dominance

### C. Social preferences

- Social dilemmas, Conditional Cooperation, Intentions

# What is a social dilemma?

A social dilemma is a situation in which individual incentives are at odds with group incentives. Specifically, individuals have the incentive to free ride, but groups do better when everyone contributes (cooperation problem).

Therefore, a necessary prerequisite is the existence of external effects (positive or negative).

- My action/consumption influences others' payoffs/consumption.

*What are examples of social dilemmas?*



# What is a social dilemma?

A social dilemma is a situation in which individual incentives are at odds with group incentives. Specifically, individuals have the incentive to free ride, but groups do better when everyone contributes (cooperation problem).

Therefore, a necessary prerequisite is the existence of external effects (positive or negative).

- My action/consumption influences others' payoffs/consumption.

*What are examples of social dilemmas?*

# Examples of social dilemmas

- Public goods (streets, bridges, parks, schools, ...)
- Common pool resources (natural resources, ...)
  
- Hunting and gathering
- Environmental protection
- Teamwork (e.g., a group project)
- Organizations
- Politics & Voting
- Charities
- ...

# Why can't a market solve the problem?

Under certain assumptions markets supply an efficient amount of *private* goods.

However, in general private markets do not supply an efficient amount of public goods.

Reason: private marginal benefit  $\neq$  social marginal benefit  
(Samuelson 1954, Lindahl 1919).

# Why can't a market solve the problem?

Under certain assumptions markets supply an efficient amount of *private* goods.

However, in general private markets do not supply an efficient amount of public goods.

Reason: private marginal benefit  $\neq$  social marginal benefit  
(Samuelson 1954, Lindahl 1919).

→ If we have to rely on private provision: inefficient undersupply of the public good.

See also:

- Gordon (1954) fishermen; Hardin (1968) pasture : “*tragedy of the commons*”
- Olson (1971): “*problem of collective action*”
- Dawes (1980): “*social dilemma*”

*Are the prospects really that bleak?*

# Prisoners' dilemma game as a simple workhorse

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	30\$ , 30\$	10\$ , 45\$
	Defect	45\$ , 10\$	15\$ , 15\$

- Binary choice social dilemma
- Game is played simultaneously and one-shot
- D/D is Nash equilibrium (one-shot game)
- C/C would be socially efficient

# Prisoners' dilemma game as a simple workhorse

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	30\$ , 30\$	10\$ , 45\$
	Defect	45\$ , 10\$	15\$ , 15\$

- Binary choice social dilemma
- Game is played simultaneously and one-shot
- D/D is Nash equilibrium (one-shot game)
- C/C would be socially efficient

**Usual finding:  
About 40 to 60%  
of subjects  
cooperate**

**BUT:** maybe oversimplified?!

# Another simple workhorse: VCM

## Voluntary contribution mechanism (VCM)

- aka „Public-good game“
- Introduced in the early 1980s by James Walker and Marc Isaac

### **Usual design in the VCM:**

- Groups with  $n$  members
- Each member has endowment of  $E$  “tokens”
- Each group member decides simultaneously about contribution  $c_i$ 
  - $c_i$  investment in public good
  - private good:  $(E - c_i)$
- Payoff function for each group member  $i$ :

$$\pi_i = (E - c_i) + \alpha \sum_{j=1}^n c_j$$

# VCM aka Public Good Game

You are playing this game together with 3 of your classmates. Each of you have been given 100 euros. Every group member has to decide how many euros he or she wishes to contribute to the common pool. Every euro contributed to the common pool gets multiplied by a factor of 2 and gets distributed evenly among group members. Your final payoff is the amount of money you were given in the beginning (EUR100), minus your contribution, plus the returns from the common pool. So for example, if everyone contributes 80% of their initial wealth then each member's final payment would be:  
 $100 - 80 + (4 * 80 * 2)/4 = 20 + 160 = \text{EUR}180.$

How much do you wish to contribute to the common pool?

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

Select a contribution



# What characteristics define a public good?

## ***non-rival (in consumption):***

- rivalry means that one's consumption precludes the consumption of another (e.g., ice cream).
- Key difference Public Goods vs. Common Pool Resources is the non-rivalry assumption.

## ***non-excludable (in access)***

- excludability means that consumers can be excluded from the benefits of the PG (e.g., pay-tv)

# What characteristics define a public good?

## ***non-rival (in consumption):***

- rivalry means that one's consumption precludes the consumption of another (e.g., ice cream).
- Key difference Public Goods vs. Common Pool Resources is the non-rivalry assumption.

## ***non-excludable (in access)***

- excludability means that consumers can be excluded from the benefits of the PG (e.g., pay-tv)

Because of PGs non-rivalry and non-excludability, they are susceptible to **free-riding**. Why?

- Non-excludability is key. People can free ride and enjoy the benefits of the resource without paying any associated costs because they cannot be kicked out of the group of beneficiaries.

# Categorizing social dilemmas

		Is consumption rival?	
		Yes	No
Can persons be excluded?	Yes	Private Good	Toll Good
	No	Common Pool Resource	Public Good (VCM)

# Categorizing social dilemmas

		Is consumption rival?	
		Yes	No
Can persons be excluded?	Yes	Private Good	Toll Good
	No	Common Pool Resource	Public Good (VCM)

# The VCM design:

$$\pi_i = (E - c_i) + \alpha \sum_{j=1}^n c_j$$

- The MPCR, marginal per capita return,  $\alpha$  is the return to an individual from making a contribution to the public account.
- Upper and lower bounds on  $\alpha$  for VCM being a public-goods problem (and social dilemma):
  - $\alpha < 1$  : otherwise the “dominant strategy” to allocate all tokens to the public good
  - $1/n < \alpha$ : otherwise it is socially efficient to keep all tokens because PG is so unproductive
    - if everyone keeps endowment, we end up with a total of  $nE$ ;
    - if everyone contributes, we end up with  $n(\alpha nE)$ . Thus, PG is beneficial for the individual
      - $\alpha nE > nE$
      - $\Leftrightarrow \alpha nE > E$
      - $\Leftrightarrow \alpha n > 1$
      - $\Leftrightarrow \alpha > 1/n$

# Usual findings in the VCM

Frequently used parameters:  $n=4$ ;  $E=20$ ;  $\alpha=.4$

Standard Prediction:  $c_i = 0, \forall i$ , which implies an inefficient level of contribution!

Do we usually observe this in experiments?

No, initial average contributions are roughly 40 to 60% of the initial endowment

# VCM aka Public Good Game

You are playing this game together with 3 of your classmates. Each of you have been given 100 euros. Every group member has to decide how many euros he or she wishes to contribute to the common pool. Every euro contributed to the common pool gets multiplied by a factor of 2 and gets distributed evenly among group members. Your final payoff is the amount of money you were given in the beginning (EUR100), minus your contribution, plus the returns from the common pool. So for example, if everyone contributes 80% of their initial wealth then each member's final payment would be:  
 $100 - 80 + (4 * 80 * 2)/4 = 20 + 160 = \text{EUR}180.$

How much do you wish to contribute to the common pool?

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

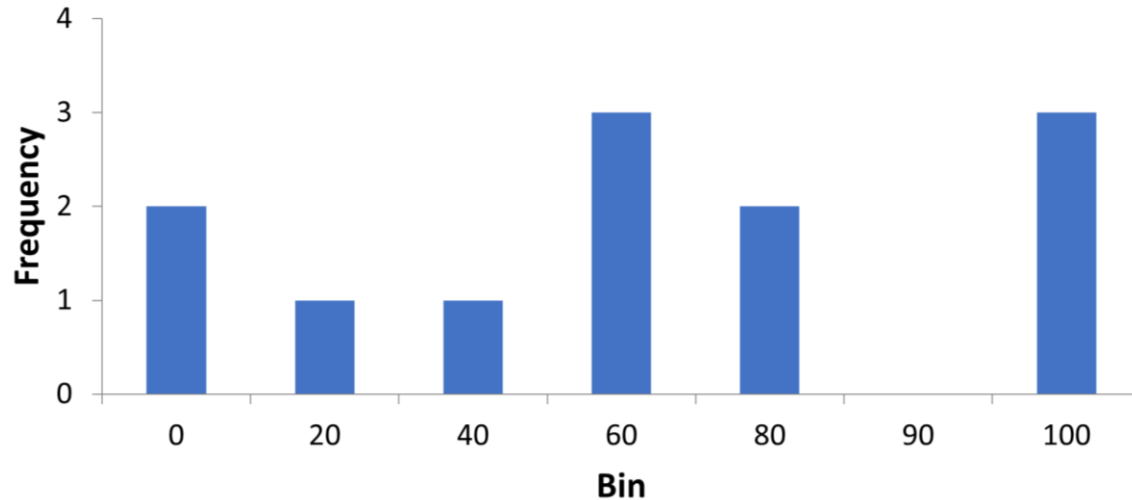
Select a contribution



# Public Good Game TUM

Typical initial average contributions are 40 to 60% of the initial endowment

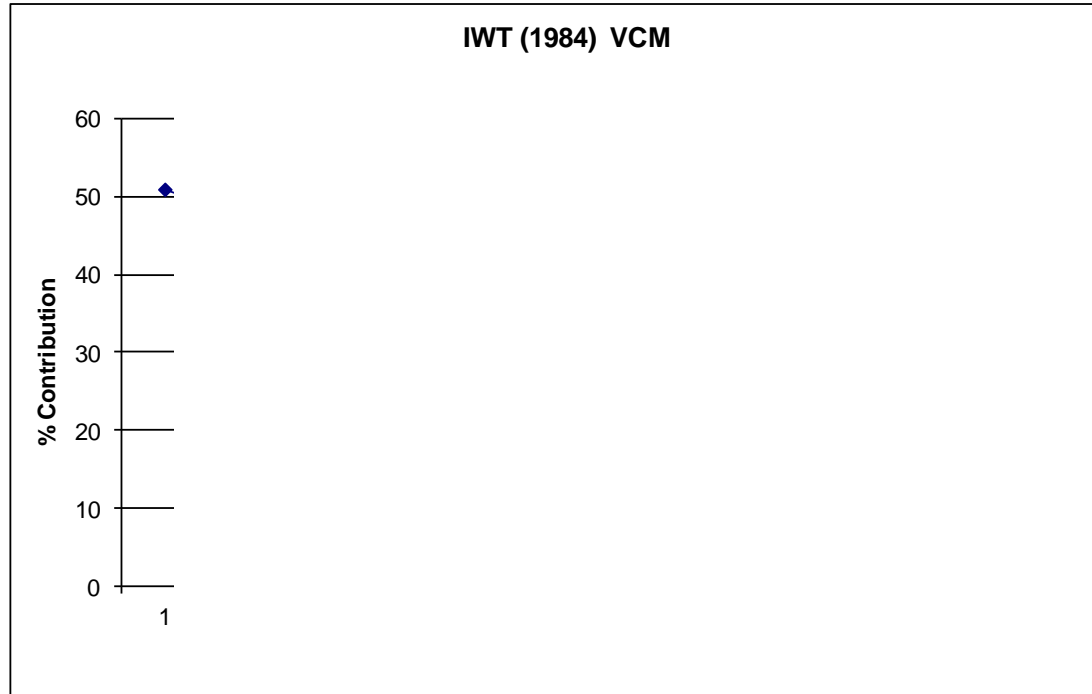
In our experiment: average contributions were 54.9%





# Results from the first VCM

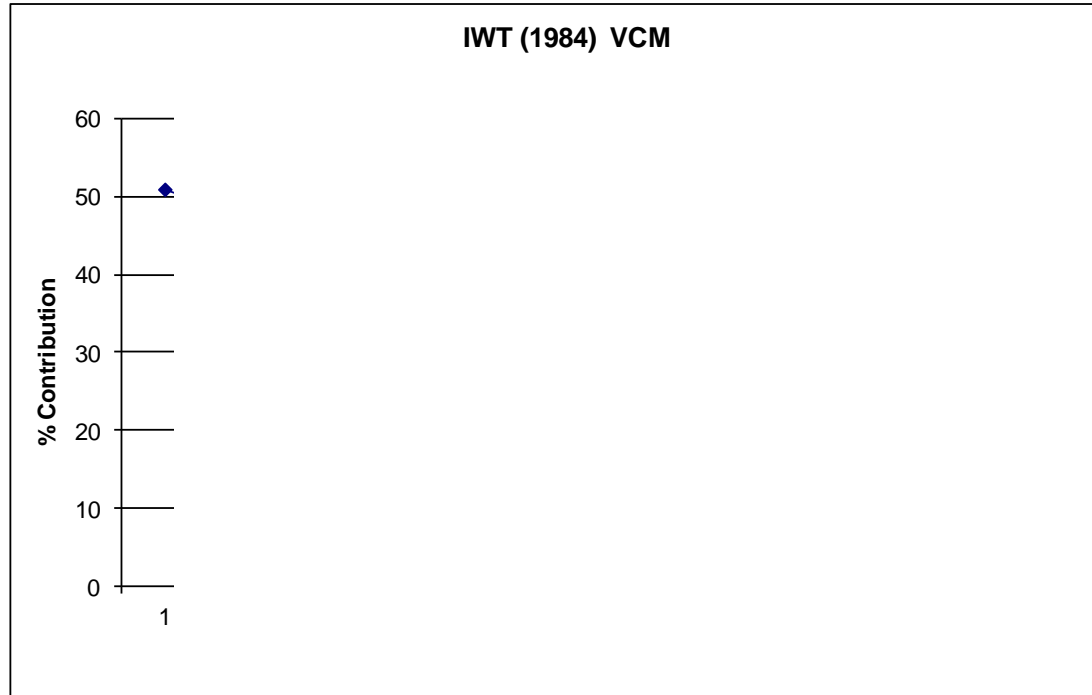
Isaac, Walker & Thomas 1984



# Results from the first VCM

Isaac, Walker & Thomas 1984

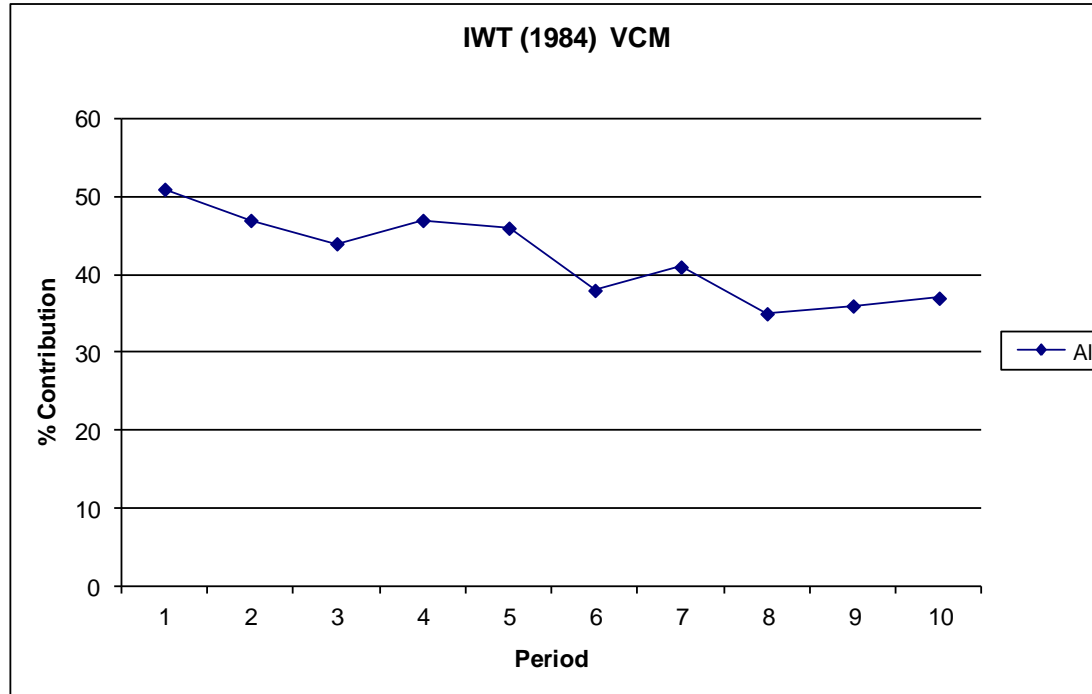
*What happens if VCM is played repeatedly? Do they reach equilibrium?*



# Results from the first VCM

Isaac, Walker & Thomas 1984

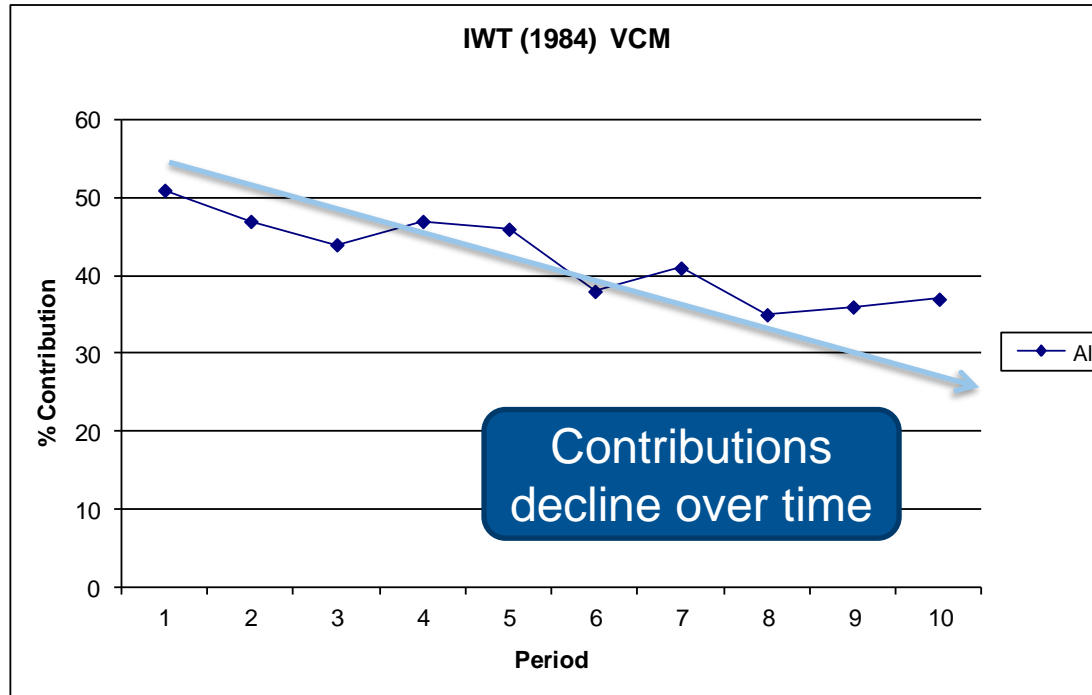
*What happens if VCM is played repeatedly? Do they reach equilibrium?*



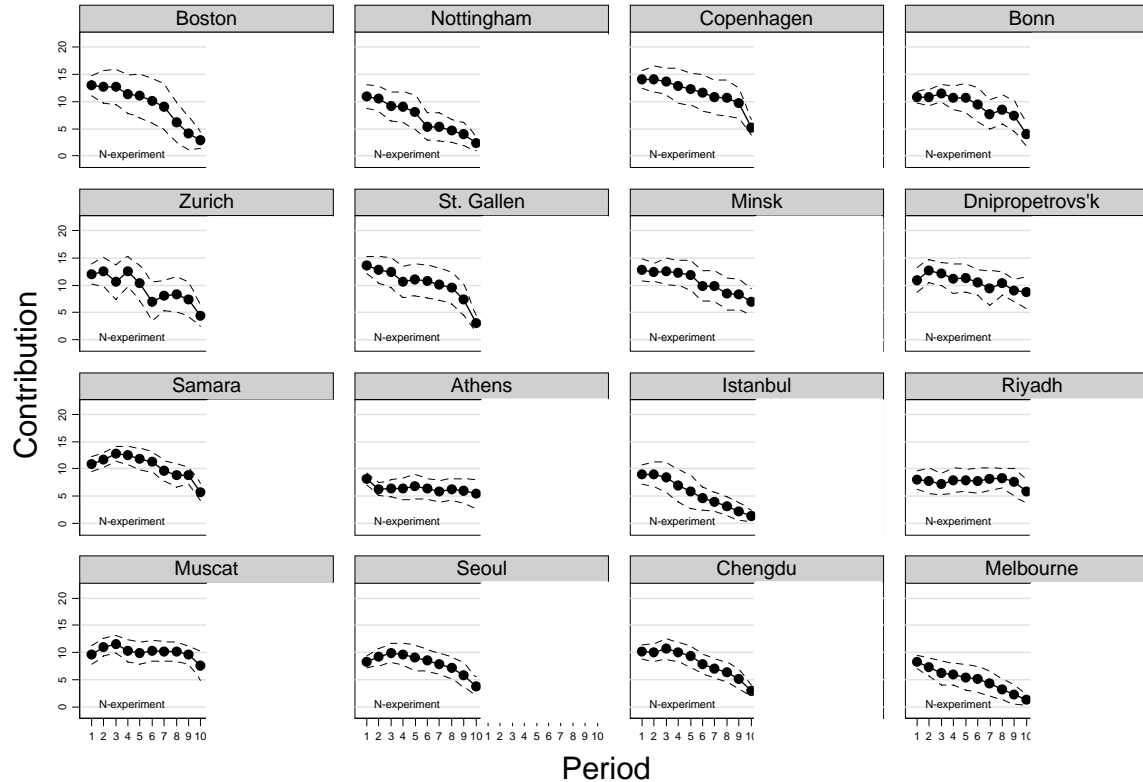
# Results from the first VCM

Isaac, Walker & Thomas 1984

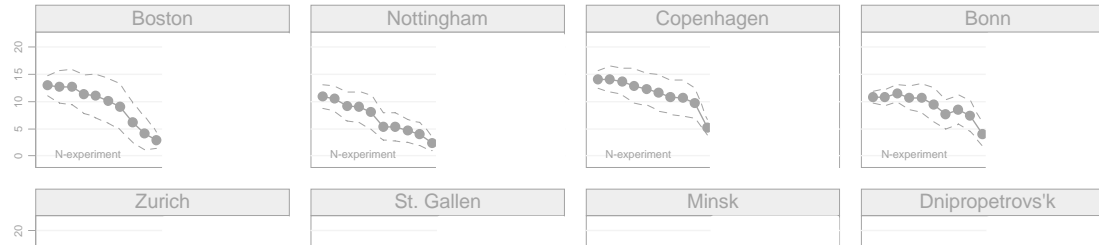
*What happens if VCM is played repeatedly? Do they reach equilibrium?*



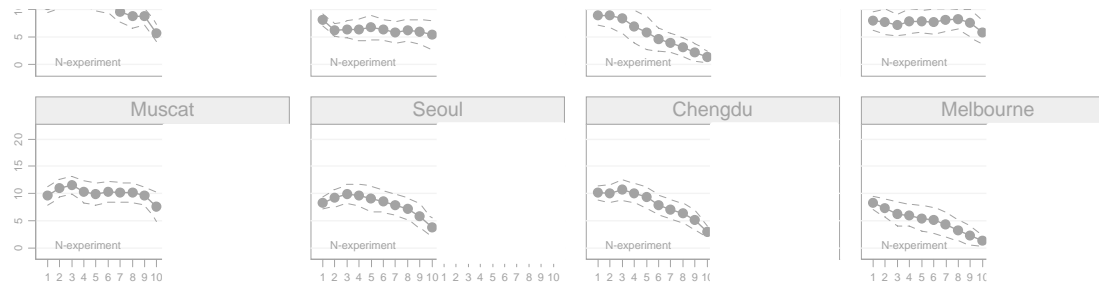
# Usual finding in the VCM



# Usual finding in the VCM



- Initial cooperation rate at roughly 40 to 60 percent (of full cooperation)
- Drops to virtually zero in later periods
- In final periods full defection is the most frequent choice



Period

# Why do we observe this contribution pattern?

We observe substantial cooperation in the beginning and a decline of cooperation over time

*Why?*

# Why do we observe this contribution pattern?

We observe substantial cooperation in the beginning and a decline of cooperation over time

*Why?*

- Maybe the economic standard assumption of self-centered, rational agents is right  
(and thus the Nash prediction of 0-cooperation)

People...

- ...just make errors (in the beginning)?
- ...just have to learn the Nash-equilibrium point?



# Why do we observe this contribution pattern?

We observe substantial cooperation in the beginning and a decline of cooperation over time

*Why?*

- Maybe the economic standard assumption of self-centered, rational agents is right  
(and thus the Nash prediction of 0-cooperation)

People...

- ...just make errors (in the beginning)?
  - ...just have to learn the Nash-equilibrium point?
- 
- Maybe the assumptions are wrong, and something else is going on in the VCM?

*How can we test this?*

# Do people “learn Nash”?

Andreoni (1988)

## Experimental Design:

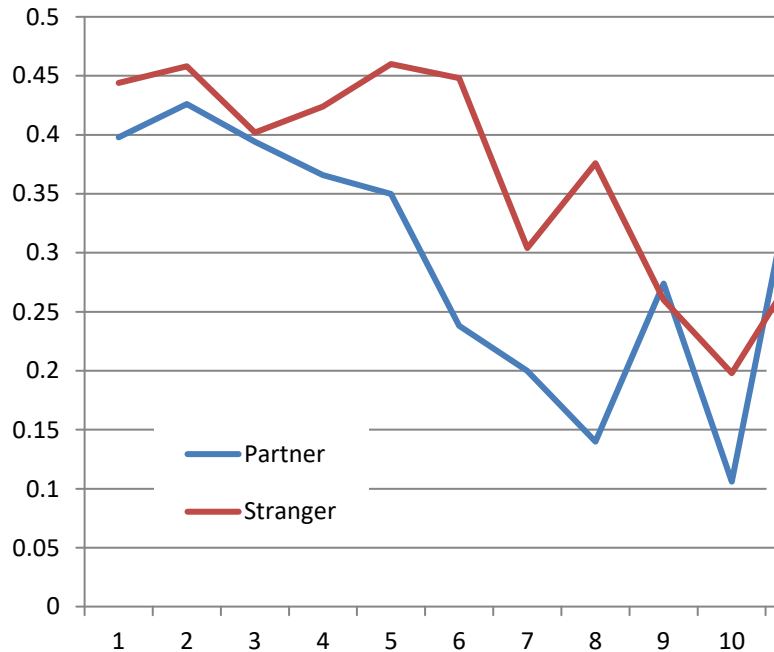
- 10 period VCM
- After period 10 surprise restart of the whole experiment

## What does this experiment measure:

- Effect of a surprise restart
- If behavior in the VCM is about learning Nash, subject's should behave similar in the period before the restart and directly after

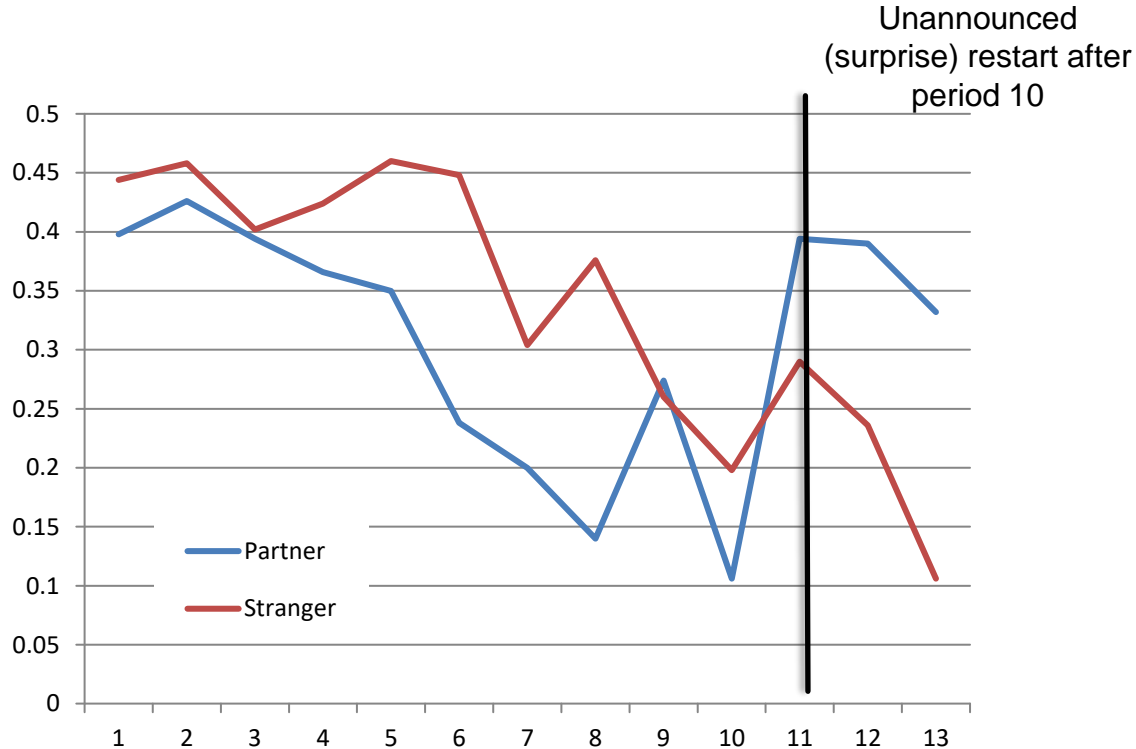
# Do people “learn Nash”?

Andreoni (1988)



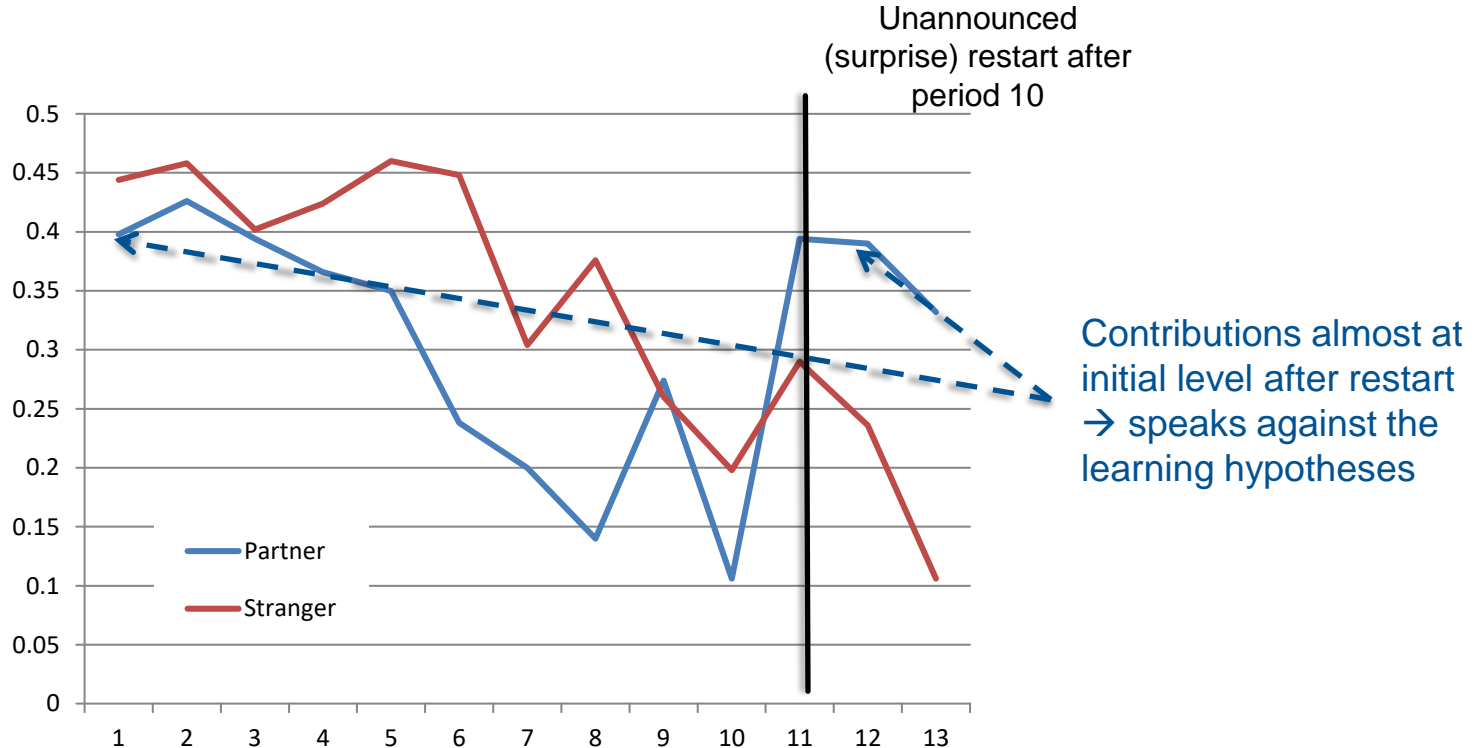
# Do people “learn Nash”?

Andreoni (1988)



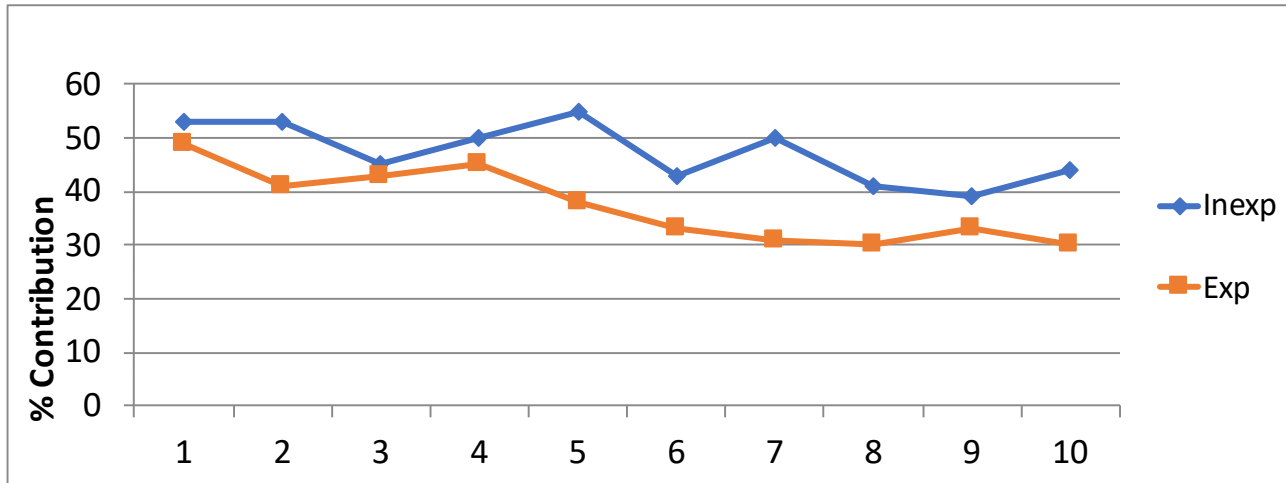
# Do people “learn Nash”?

Andreoni (1988)



# Do people “learn Nash”?

Isaac, Walker and Thomas (1984)



Experience in the sense of having participated in a VCM before has only a slight level effect, but again we observe a restart effect

# Why then do people cooperate?

Maybe not all subjects are self-centered, money-maximizing, rational agents?

- Maybe people care about outcomes and inequality?
  - Fehr and Schmidt (1999, QJE) argue that people have an inequality aversion (advantageous as well as disadvantageous).

$$u_{(\alpha_i, \beta_i)}(x_i, x_j) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}$$

- $\alpha \geq 0$  (“envy”),  $\beta \geq 0$  (“guilt”)

# Fehr-Schmidt-Preferences, example Prisoner's Dilemma

$$u_{(\alpha_i, \beta_i)}(x_i, x_j) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}$$

- $\alpha \geq 0$  ("envy"),  $1 \geq \beta \geq 0$  ("guilt")

$i / j$	$C$	$D$
$C$	$a, a$	$c, b$
$D$	$b, c$	$d, d$



$i / j$	$C$	$D$
$C$	$a, a$	$c - \alpha_i(b - c), b - \beta_j(b - c)$
$D$	$b - \beta_i(b - c), c - \alpha_j(b - c)$	$d, d$



# Fehr-Schmidt-Preferences, example Prisoner's Dilemma

$$u_{(\alpha_i, \beta_i)}(x_i, x_j) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}$$

- $\alpha \geq 0$  ("envy"),  $1 \geq \beta \geq 0$  ("guilt")

<i>i / j</i>	<i>C</i>	<i>D</i>
<i>C</i>	<i>a, a</i>	<i>c, b</i>
<i>D</i>	<i>b, c</i>	<i>d, d</i>

→

<i>i / j</i>	<i>C</i>	<i>D</i>
<i>C</i>	<i>a, a</i>	$c - \alpha_i(b - c), b - \beta_j(b - c)$
<i>D</i>	$b - \beta_i(b - c), c - \alpha_j(b - c)$	<i>d, d</i>

- Example:  $\alpha = 0.8$ ,  $\beta = 0.5$

**P 2**

	<b>C</b>	<b>D</b>
<b>P 1</b>	C	D
	30\$, 30\$	10\$, 45\$
	45\$, 10\$	15\$, 15\$

**P 2**

	<b>C</b>	<b>D</b>
<b>P 1</b>	C	D
	30\$, 30\$	-18\$, 27.5\$
	27.5\$, -18\$	15\$, 15\$

# Why then do people cooperate?

Maybe not all subjects are self-centered, money-maximizing, rational agents?

- Maybe people care about outcomes and inequality?
  - Fehr and Schmidt (1999, QJE) argue that people have an inequality aversion (advantageous as well as disadvantageous).

$$u_{(\alpha_i, \beta_i)}(x_i, x_j) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}$$

- $\alpha \geq 0$  (“envy”),  $\beta \geq 0$  (“guilt”)
- Maybe people just “feel good” when cooperating?
  - Palfrey and Prisbrey (1997, AER) argue that people receive a “warm glow” from cooperating
  - they cooperate *unconditionally*, i.e., independent of what other group members do (similar to altruists)
  - For some time, this used to be the leading explanation why people cooperate

# Why then do people cooperate? Why then do contributions decline?

But why the decline over time? Does it have to do something with intentions and reciprocity?

Today, most common explanation is existence of **conditional cooperators**:

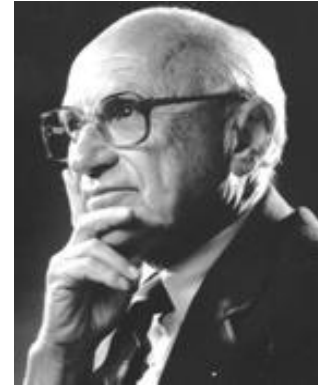
- People “get it” from the start, but the only way that cooperative types can get back at free-riders in the VCM is to withhold their own contributions.

# Conditional cooperation is intuitive

"... we might all of us be willing to contribute to the relief of poverty, provided everyone else did. We might not be willing to contribute the same amount without such assurance."

Milton Friedman

*Capitalism and Freedom*, 1962, p.191



# How to identify conditional cooperation?

Let's first look again at the binary situation, the prisoners' dilemma game

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	30\$ , 30\$	10\$ , 45\$
	Defect	45\$ , 10\$	15\$ , 15\$

Problem:

- if we see someone choosing C, is it an altruist or a conditional cooperator?
- If we see someone choosing D, is it a free-rider or a conditional cooperator?

# How to identify conditional cooperation?

Let's first look again at the binary situation, the prisoners' dilemma game

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	30\$ , 30\$	10\$ , 45\$
	Defect	45\$ , 10\$	15\$ , 15\$

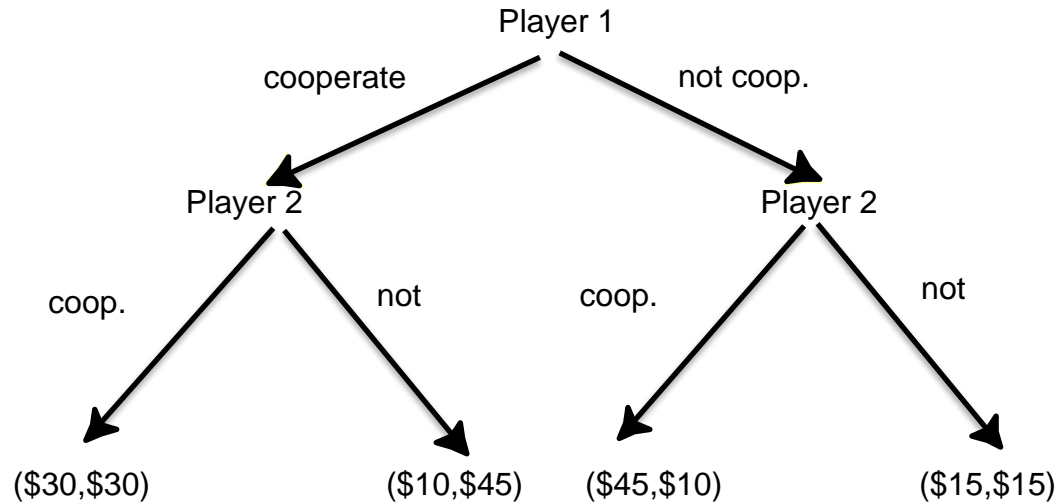
Problem:

- if we see someone choosing C, is it an altruist or a conditional cooperators?
- If we see someone choosing D, is it a free-rider or a conditional cooperators?

Answer depends on the **beliefs**, but how to elicit truthfully?

# What if the game is played sequentially?

Fehr, Kosfeld, and Weibull (2003) or Burks, Carpenter, and Goette (2005)



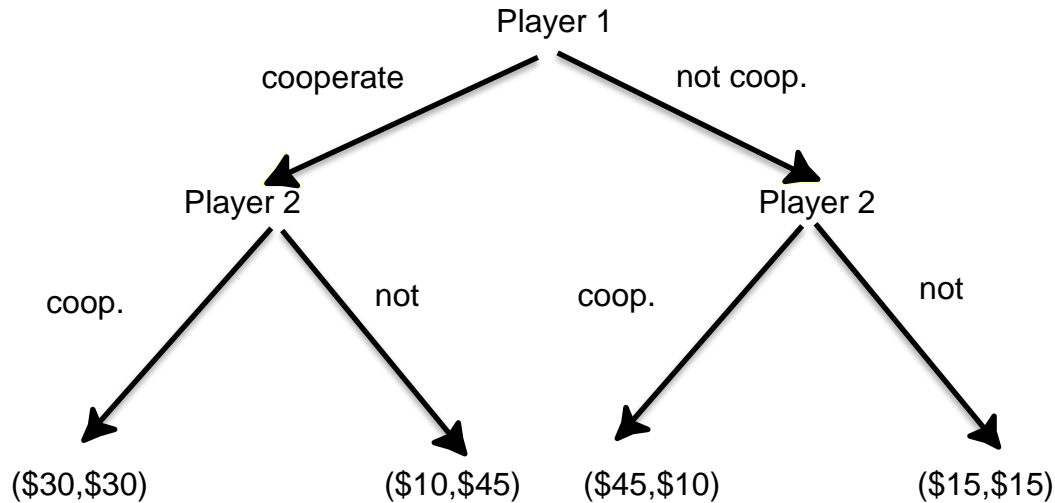
# What if the game is played sequentially?

Fehr, Kosfeld, and Weibull (2003) or Burks, Carpenter, and Goette (2005)

*What should a free-rider do?*

*What should an altruist do?*

*What should a conditional cooperator do?*





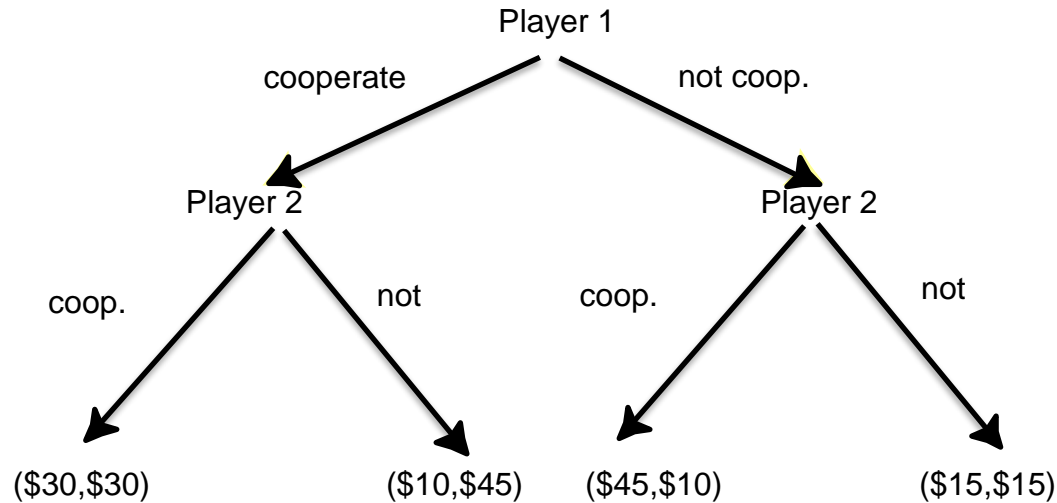
# What if the game is played sequentially?

Fehr, Kosfeld, and Weibull (2003) or Burks, Carpenter, and Goette (2005)

*What should a free-rider do?*

*What should an altruist do?*

*What should a conditional cooperator do?*



Trick is to use the “**strategy method**” (Selten 1967): Play as 2 in both branches

# How do students play the sequential PD?

Fehr, Kosfeld and Weibull (2003)

## Experimental Design:

- 96 Swiss students play both roles in the Sequential PD.
- In player 1 role, they decide to cooperate or to not cooperate
- In player 2 role, they submit a strategy: (if 1 cooperates, if 1 does not cooperate).

## What does this experiment measure:

- Use these strategies to “type” the players as
  - selfish
  - reciprocal (conditional cooperators)
  - altruistic
  - anti-reciprocal

# How do students play the sequential PD?

Fehr, Kosfeld and Weibull (2003)

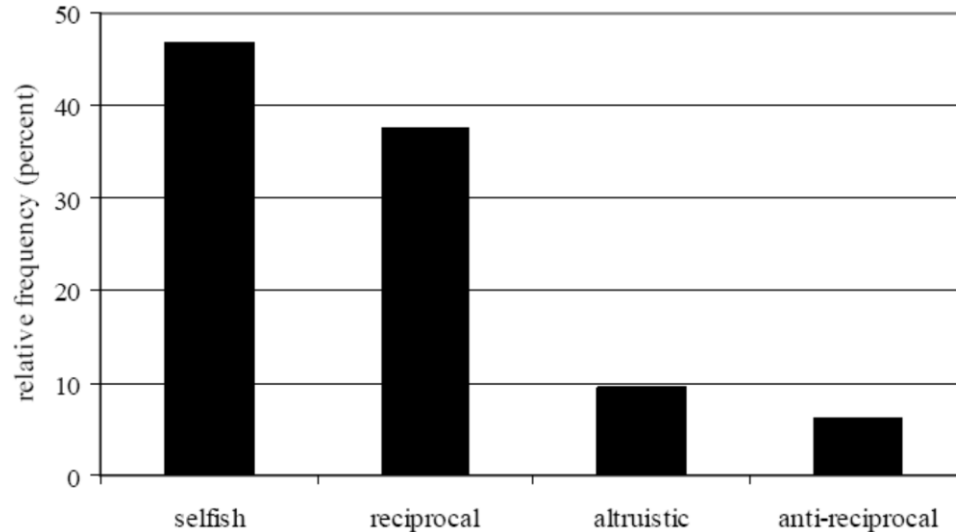


Figure 2: Distribution of revealed PD preferences

# How do students play the sequential PD?

Fehr, Kosfeld and Weibull (2003)

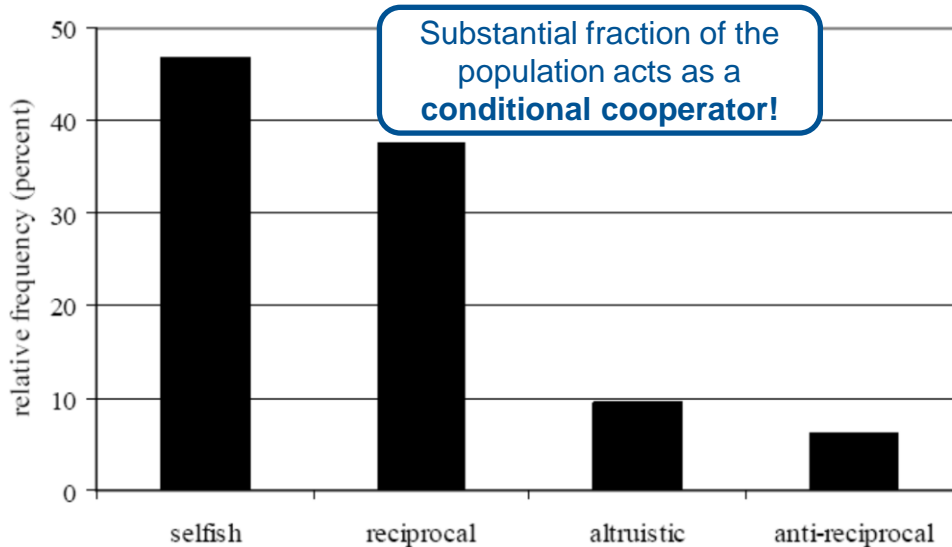


Figure 2: Distribution of revealed PD preferences  
Cond.  
Coop.

# Use the same classification approach in the VCM?

## ***'conditional cooperators'*** in the VCM:

- people who are willing to contribute more to a public good the more others contribute  
→ maybe more precisely: **conditional contributors**

# Use the same classification approach in the VCM?

## ***‘conditional cooperators’ in the VCM:***

- people who are willing to contribute more to a public good the more others contribute
- maybe more precisely: **conditional contributors**

## **Indirectly: look at patterns of behavior (problematic)**

- Keser, van Winden Scand J Economics 2000 (partner > stranger)
- Sonnemans, Schram, Offerman, EconLetters 1999 (replacement)

## **Directly: apply a variant of the ‘strategy method’ (Selten, 1967) to elicit subjects’ type**

- Fischbacher, Gächter, Fehr (Economics Letters 2001)
- Falk & Fischbacher (EER 2002)

# Conditional Cooperators

Fischbacher, Gächter, Fehr (2001)

## Experimental Design:

- Groups with n=4
- Three-stage VCM:

$$\pi_i = 20 - c_i + 0.4 \sum_{j=1}^4 c_j$$

- **Stage 1: Unconditional contribution**
- **Stage 2: Conditional contribution**

Subjects indicate for each average contribution level of other group members how much they want to contribute to the public good.

- **Stage 3: One person of the group is randomly selected**  
→ conditional contribution (on the average of the other three subjects not selected)

# Conditional Cooperators

Fischbacher, Gächter, Fehr (2001)

Periode 1 von 1 Verbleibende Zeit [sec]: 28

Ihr bedingter Beitrag zum Projekt (Beitragstabelle)

0	<input type="text"/>	7	<input type="text"/>	14	<input type="text"/>
1	<input type="text"/>	8	<input type="text"/>	15	<input type="text"/>
2	<input type="text"/>	9	<input type="text"/>	16	<input type="text"/>
3	<input type="text"/>	10	<input type="text"/>	17	<input type="text"/>
4	<input type="text"/>	11	<input type="text"/>	18	<input type="text"/>
5	<input type="text"/>	12	<input type="text"/>	19	<input type="text"/>
6	<input type="text"/>	13	<input type="text"/>	20	<input type="text"/>

Hilfe  
Geben Sie in den Feldern ein, welchen Beitrag zum Projekt Sie leisten, wenn die anderen im Durchschnitt den Beitrag zum Projekt geleistet haben, der links vom Eingabefeld steht.  
Wenn Sie alles eingegeben haben, drücken Sie "OK".



# Conditional Cooperators

Fischbacher, Gächter and Fehr (2001)

## How are types defined?

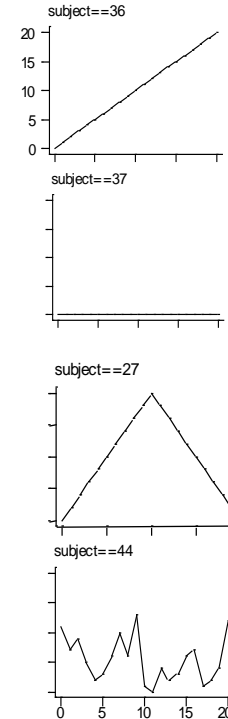
Conditional Cooperators.: Spearman's  $\rho > 0$ , p-value  $< 0.001$

Free-riders: trivial (always give 0)

Hump-shaped: increasing up to some point, then decreasing

Others: everyone else, e.g.,

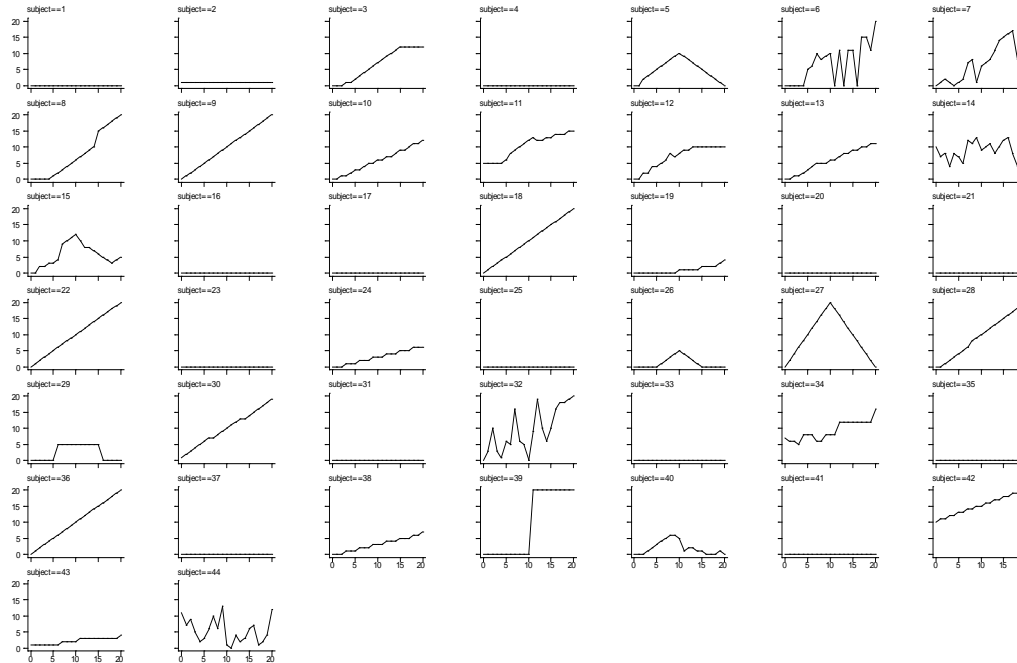
Prof. Dr. Sebastian Goerg & Dr. Orestis Kopsacheilis | Behavioral Economics



# Conditional Cooperators

Fischbacher, Gächter and Fehr (2001)

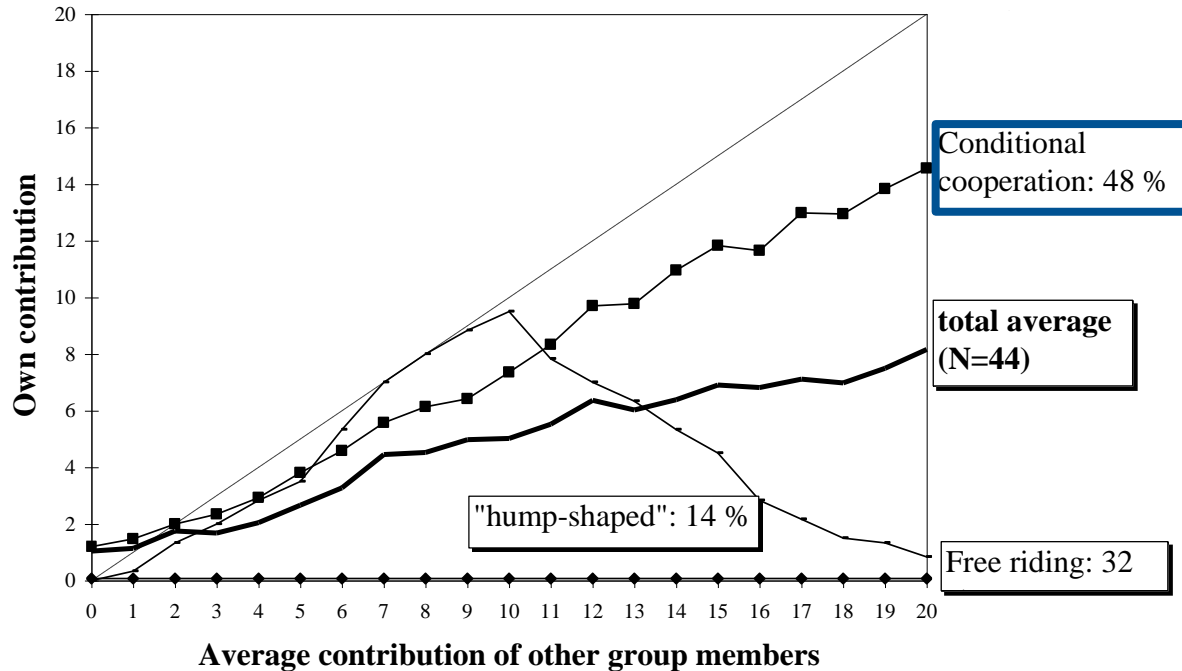
**Finding: Subjects are heterogeneous**



# Conditional Cooperators

Fischbacher, Gächter and Fehr (2001)

**Finding: Substantial fraction of the population acts as a conditional cooperator!**



# Findings provide a potential explanation for the unraveling of cooperation.

Many people are willing to cooperate conditional on others' cooperation.

A large minority of the subjects free-rides fully irrespective of what others do.

The free-riders induce the conditional cooperators to cease to cooperate once they realize that there are free-riders in the group. Potentially explains the decay in cooperation over time.

*Can we test this hypotheses?*

# Common Pool Resource Game (CPR)

Up to now, the problem of voluntary provision of public goods was considered (VCM)  
But what if consumption is rival, i.e., when we have negative externalities?

		Is consumption rival?	
		Yes	No
Can persons be excluded?	Yes	Private Good	Toll Good
	No	Common Pool Resource	Public Good (VCM)

# Common Pool Resource Game (CPR)

When thinking of a public goods, some people think of it as a common resource

- tragedy of the commons (“Allmende Problem”)
- Examples: fishing, foresting

The CPR models extraction from a resource that reduces the others’ possibility of extraction

- Here consumption is rival and access is non-excludable.
- In other words, my extraction creates a negative externality for you.
- Given this, people cooperate by limiting their extraction from the CPR.

# An example of a CPR experiment

Cardenas, Stranlund, and Willis (2000)

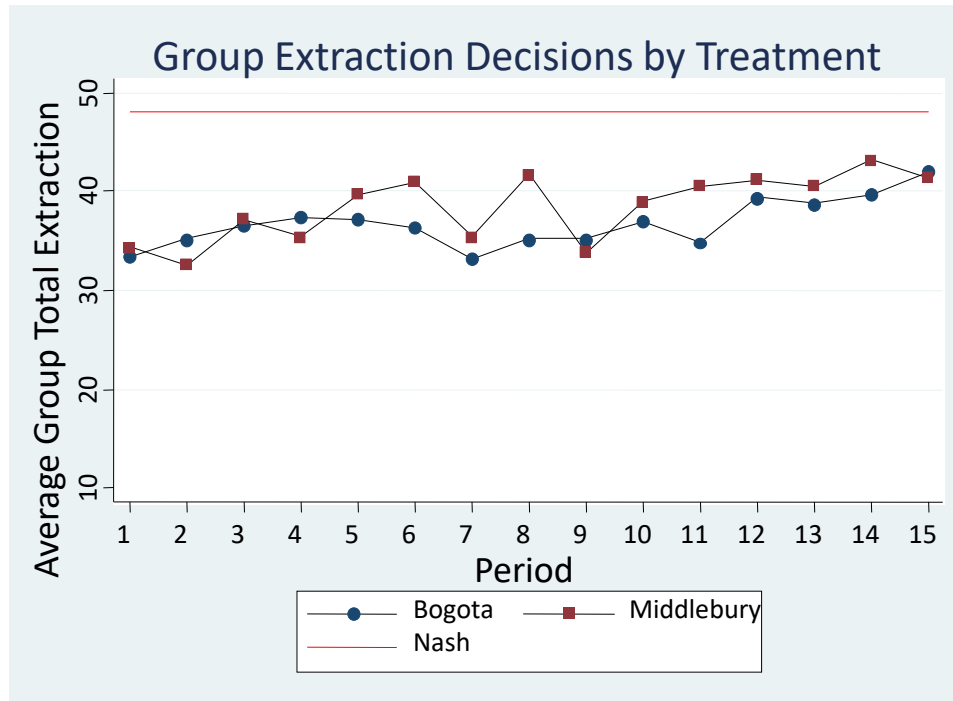
## Design:

- Columbia: rural population with significant interests in local natural resources and environmental quality
- Payoffs for the game:
  - generated from a model of individual efforts to collect firewood from local forests.
  - Wage in formal sector, benefit from collecting firewood, water quality and firewood collection, valuation of water quality
  - Private and social interests diverge (higher levels of firewood extraction heighten soil erosion and ultimately damage local water quality)
- Framing: subjects are fully aware they are playing a game with this specific relationship between firewood extraction and water quality in place
- Groups of 8 play the CPR
- The underlying payoff function was presented (for obvious reasons) in form of a payoff table

# An example of a CPR experiment

Cardenas, Stranlund, and Willis (2000)

**Finding: Behavior in the CPR closely resembles what is observed in the VCM games**

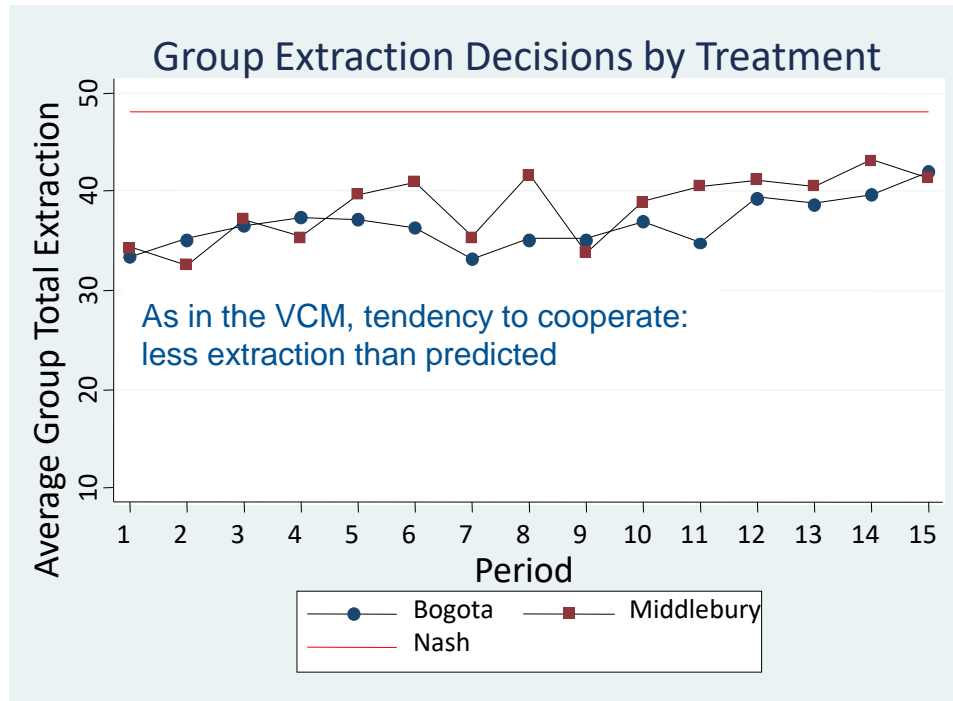




# An example of a CPR experiment

Cardenas, Stranlund, and Willis (2000)

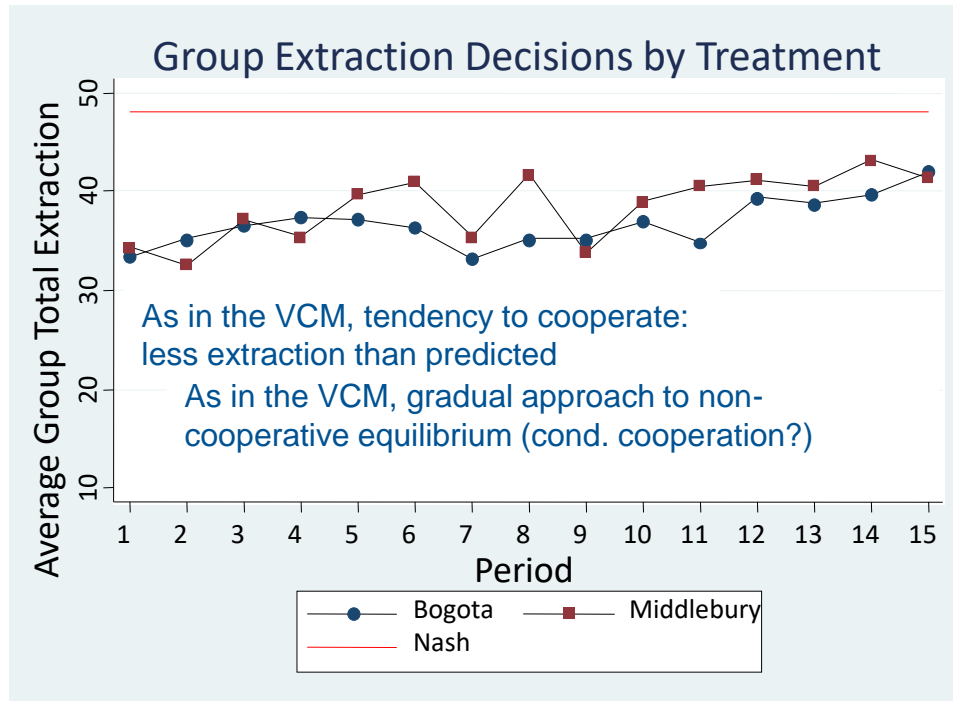
**Finding: Behavior in the CPR closely resembles what is observed in the VCM games**



# An example of a CPR experiment

Cardenas, Stranlund, and Willis (2000)

**Finding: Behavior in the CPR closely resembles what is observed in the VCM games**



# Often external regulations are implemented to solve such problems

External intervention:

- an outside body (e.g., the state) monitors and sanctions asocial behavior, e.g.
  - Controls and fines to hinder exploitation of resources
  - (tax schemes to provide public goods)

As we know from everyday life, this often solves the social dilemma

- This is particularly true for deterrent sanctions...
- ...but it must not necessarily be the case....

# The problems with external intervention

Ostrom (1990), reviewing many case studies suggests that external government intervention in the management of CPRs can just as often harm the CPR as help it.

In some cases, the government has less information than local users and this lack of information is harmful. In other cases, formal exogenous institutions (like monitoring and fines) “crowd out” whatever pro-social tendencies that local users may have.

*What about endogenous interventions?*



Elinor Ostrom  
(1933-2012)

# Endogenous interventions

## Endogenous regulation (self-governance)

- the members of the group establish rules and norms for dealing with free riders
- they impose sanctions if norms are broken.

But: endogenous regulations are again a public good; now a 2<sup>nd</sup>-order public good.

To see what that means, consider the following game that tries to capture endogenous regulations...

# Decentralized punishment game

## Typical design:

Stage 1: regular VCM:

$$\pi_i = (20 - c_i) + 0.4 \sum_{j=1}^n c_j$$

### Stage 2:

- Players observe (anonymously) how much the others contributed.
- They then decide simultaneously whether to assign punishment points  $p_{ij}$  to other players.
- Each punishment point costs 1 token for the punisher.
- Each punishment point reduces the payoff of the punishee by 3 tokens.

$$\pi_i = (20 - c_i) + 0.4 \sum_{j=1}^n c_j - 1 \sum_{i \neq j} p_{ij} - 3 \sum_{j \neq i} p_{ji}$$

# Punishment as a 2<sup>nd</sup>-order public good

Payoff function is

$$\pi_i = (20 - c_i) + 0,4 \sum_{j=1}^n c_j - \sum_{i \neq j} p_{ij} - 3 \sum_{j \neq i} p_{ji}$$

## Using backward induction, what is the SP equilibrium?

- Do not punish on stage 2, because it reduces your payoff.
- Therefore, on stage 1 do not expect others to punish you for free-riding on stage 2
- Thus feel safe and do not contribute.
- Everybody would be better off if free-rider would be punished on stage 2 so that they do cooperate on stage 1 – but everybody would want the other to carry out the costly punishment → 2<sup>nd</sup>-order public good

# Punishment as a 2<sup>nd</sup>-order public good

Payoff function is

$$\pi_i = (20 - c_i) + 0,4 \sum_{j=1}^n c_j - \sum_{i \neq j} p_{ij} - 3 \sum_{j \neq i} p_{ji}$$

## Using backward induction, what is the SP equilibrium?

- Do not punish on stage 2, because it reduces your payoff.
- Therefore, on stage 1 do not expect others to punish you for free-riding on stage 2
- Thus feel safe and do not contribute.
- Everybody would be better off if free-rider would be punished on stage 2 so that they do cooperate on stage 1 – but everybody would want the other to carry out the costly punishment → 2<sup>nd</sup>-order public good

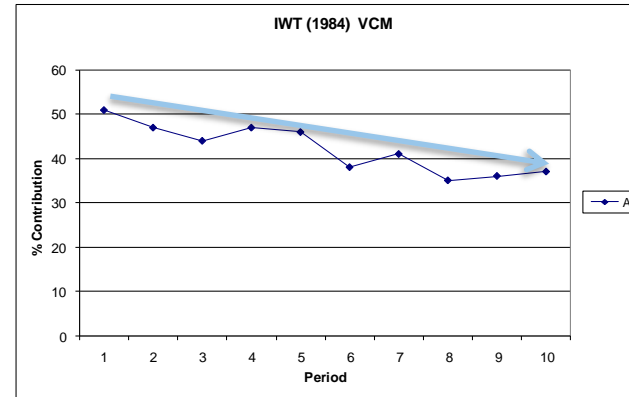
*Regardless of the theoretical solution, how do you think behavior differs under the new rules?*



# Selfish and reciprocal types in the VCM+PUN

Recall the data from a standard VCM →

Also recall that the leading hypothesis about what is going on is **conditional cooperation**.



Given conditional cooperation is the first impulse of most of the participants, punishment might provide a new means so that people don't feel the need to withhold contributions in order to get back at free riders!

# Cooperation and punishment

Fehr and Gächter (2000)

One of the first to run VCMs with endogenous regulation.

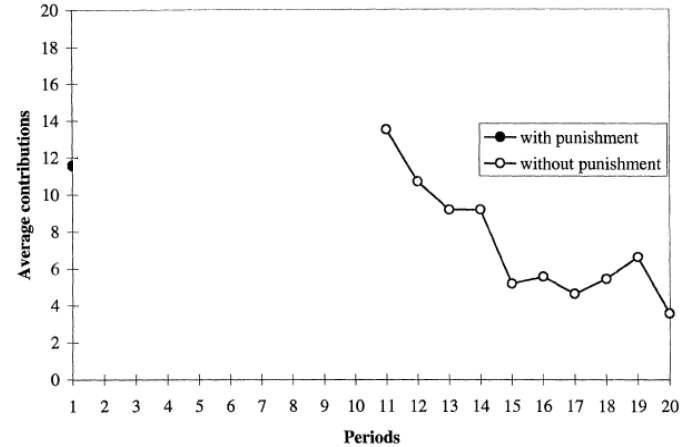
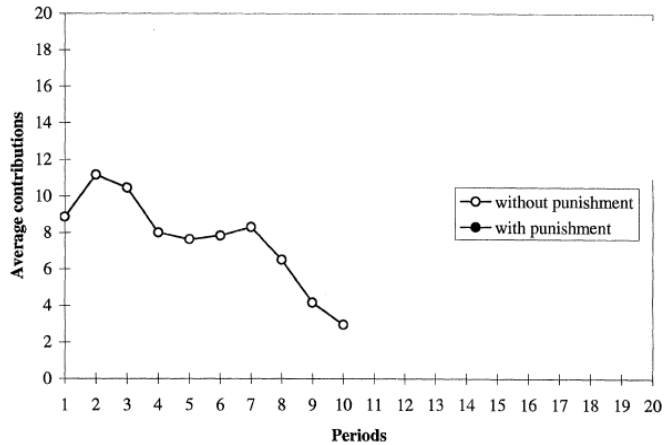
## Design:

- Payoff structure was a little more complicated
- (pay a cost to reduce the target's earnings by 10%; roughly 1:3 relation)
- However, their design is nice because they allow for within-session exposure to both the baseline (VCM) and the punishment game.
- Their design is balanced to control for potential **order-effects**: NoPun-Pun & Pun-NoPun.

# Cooperation and punishment

Fehr and Gächter (2000)

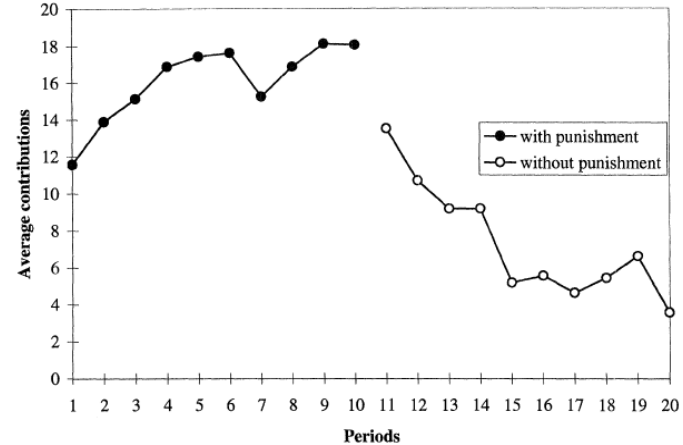
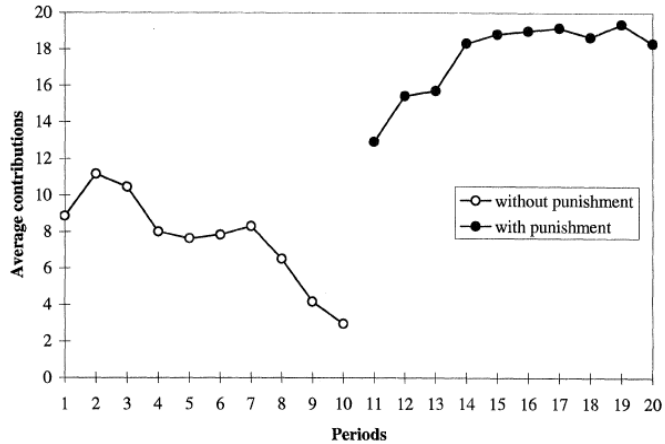
Partner condition, both orders:



# Cooperation and punishment

Fehr and Gächter (2000)

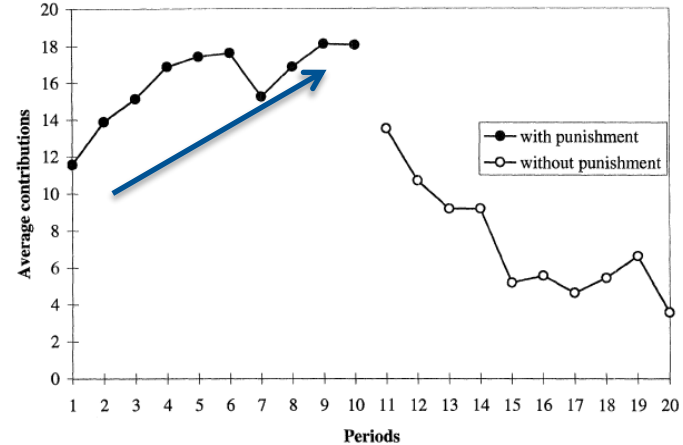
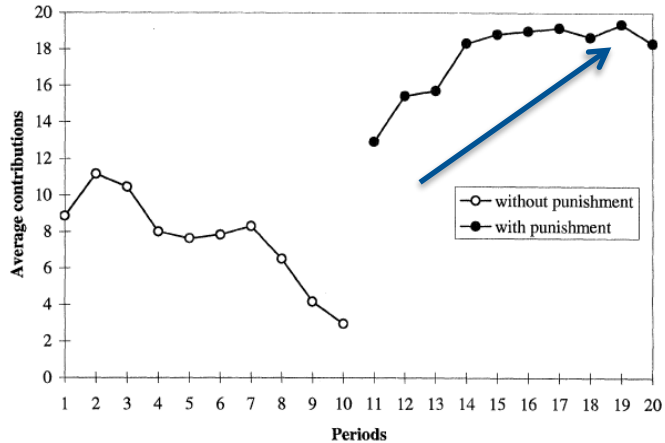
Partner condition, both orders:



# Cooperation and punishment

Fehr and Gächter (2000)

Partner condition, both orders:

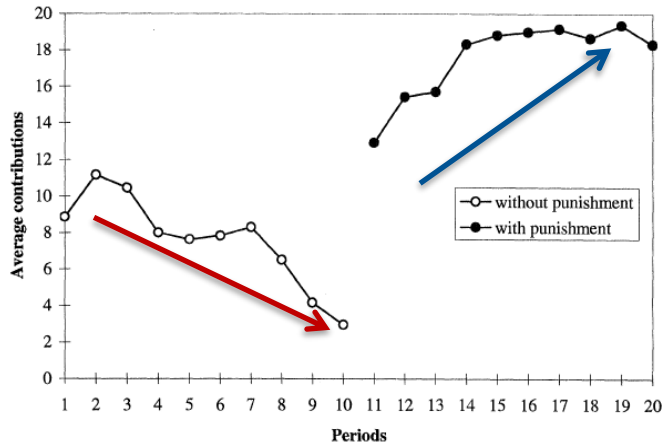


The imposition of  
punishment  
increases contributions

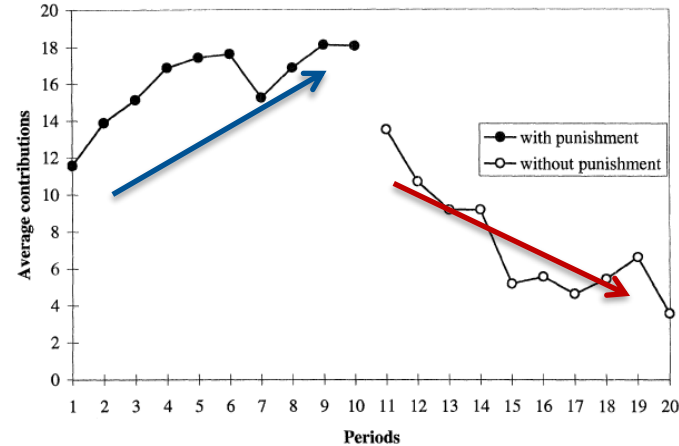
# Cooperation and punishment

Fehr and Gächter (2000)

Partner condition, both orders:



The imposition of  
punishment  
increases contributions



The removal of punishment  
allows contributions to fade.

# Decentralized punishment is very effective

- Punishment is very frequent.
- The less a player contributes the more he is punished.
- While cooperation declines without a punishment opportunity, cooperation is stable or increases with a punishment opportunity.
- Reciprocal players effectively discipline free-riders.
- 82.5% of the subjects contribute the whole endowment in the final period of the partner treatment when there is a punishment option while the majority fully defects in the final period when there is no punishment option.

# Enforcement of norms

The fact that subjects are willing to sanction free-riding behavior has important consequences for the enforcement of norms and incomplete contracts

In some sense, these informal sanctions are part of a society's social capital

Whether this is beneficial depends on the content of a norm:

- Team incentives work better than expected (cooperation norm = work hard)
- Tournament incentives work less well because cooperation means here to work less
- Less littering, less crime, tougher strikes of workers, more voting etc.



# Some (cautionary) words about social punishment

Decentralized (social) punishment is frequently effective, but:

- Strong informational assumptions (you need to know/observe what the others did)
- (Monetary) loss to society due to punishment costs (at least in the short-run)
- Might backfire, because no one guarantees that punishment is targeted at free-riders
- If the norm is susceptible to personal interpretation (for example in a complex environment), punishment can induce counter-punishment.

# VCM+PUN also effective in other cultures?

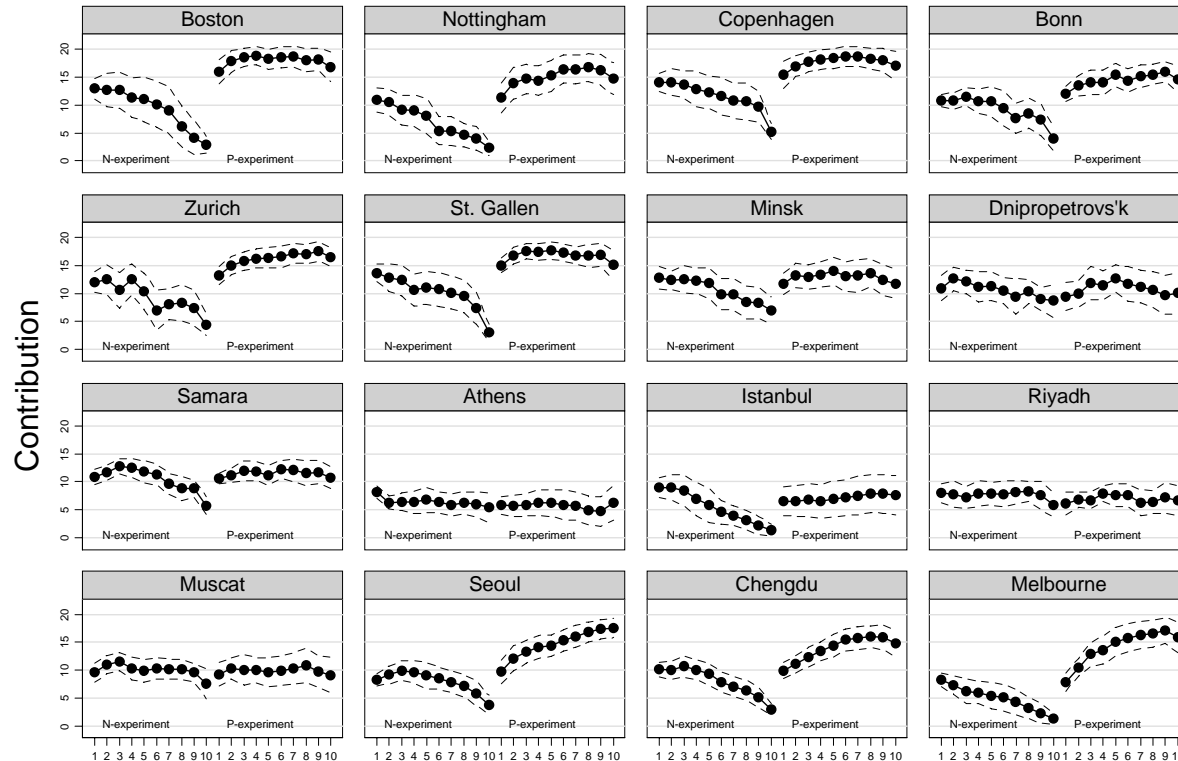
Hermann, Thöni, and Gächter (2008)

## Design:

- VCM with and without punishment in 15 countries
- 1.168 subjects
- Subjects are comparable (undergraduate students)
  
- But subject pools come from societies that differ strongly according to:
  - Trust and norms of civic cooperation (World Value Survey)
  - Cultural dimensions (Hofstede)
  - Value orientations (Inglehart)
  - Rule of Law
  - GDP per capita

# VCM+PUN also effective in other cultures?

Hermann, Thöni, and Gächter (2008)

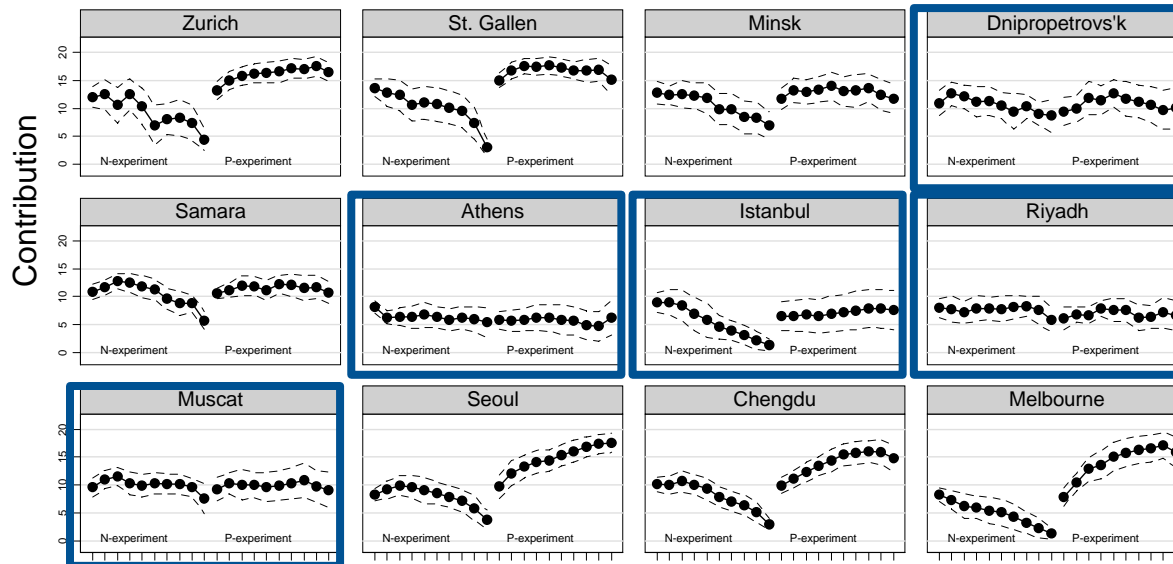


# VCM+PUN also effective in other cultures?

Hermann, Thöni, and Gächter (2008)

## Findings:

- Decentralized punishment almost always increases contributions
- But there are instances where it does not work (anti-social punishment)



# „Anti-social punishment“

Punishment of cooperators (“anti-social punishment”):

- Revenge
- Taste for increasing payoff differences
- “normative conformity” – all deviations get punished
- Dislike of “do-gooders”

*Why does punishment fail in some societies?*

# Is punishment correlated with societal-level social norms of cooperation?

Cooperation is the collectively optimal behavior. The stronger social norms of cooperation are, the more free riding is viewed as illegitimate and cooperation as desirable.

## Hypotheses:

- The stronger norms of cooperation are, the stronger should punishment of free riders be.
- The stronger norms of cooperation are, the less cooperators should be punished, because they act in the normatively desirable way.

## Measurement:

- World Values Survey: Questions on how justifiable are tax evasion, welfare fraud, and free riding on public transport.
- Theoretical scale: 10=never justified, 1=fully justified.

# Is punishment correlated with societal-level social norms of cooperation?

Cooperation is the collectively optimal behavior. The stronger social norms of cooperation are, the more free riding is viewed as illegitimate and cooperation as desirable.

## Hypotheses:

- The stronger norms of cooperation are, the stronger should punishment of free riders be.
- The stronger norms of cooperation are, the less cooperators should be punished, because they act in the normatively desirable way.



## Measurement:

- World Values Survey: Questions on how justifiable are tax evasion, welfare fraud, and free riding on public transport.
- Theoretical scale: 10=never justified, 1=fully justified.

*Hypotheses are confirmed.*

# Is punishment correlated with the strength of the Rule of Law in a society?

The rule of law indicator (World Bank) is based on a host of different variables that measure *“the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, the police, and the courts, as well as the likelihood of crime and violence”*.

- **Strong Rule of Law:** Punishment is widely perceived as being impartial (“Gewaltmonopol”) and bound by the law. Punishment is done by legitimate law enforcement institutions. Revenge is shunned.
- **Weak Rule of Law:** punishment is not impartial, is often “unfair” and arbitrary.
- Rule of Law might influence beliefs about fairness and norms of punishment in society.

## Hypothesis:

- The stronger the Rule of Law in a society, the lower should anti-social punishment be.

## Measurement:

- The rule of law indicator (World Bank)
- The theoretical range is  $-2.5$  (very weak rule of law) to  $2.5$  (very strong rule of law).




# Is punishment correlated with the strength of the Rule of Law in a society?

The rule of law indicator (World Bank) is based on a host of different variables that measure *“the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, the police, and the courts, as well as the likelihood of crime and violence”*.

- **Strong Rule of Law:** Punishment is widely perceived as being impartial (“Gewaltmonopol”) and bound by the law. Punishment is done by legitimate law enforcement institutions. Revenge is shunned.
- **Weak Rule of Law:** punishment is not impartial, is often “unfair” and arbitrary.
- Rule of Law might influence beliefs about fairness and norms of punishment in society.

## Hypothesis:

- The stronger the Rule of Law in a society, the lower should anti-social punishment be. 

## Measurement:

- The rule of law indicator (World Bank)
- The theoretical range is  $-2.5$  (very weak rule of law) to  $2.5$  (very strong rule of law).

*Hypothesis is confirmed.*

# Is punishment always a panacea?



The Hatfield Clan  
(1897)

The Hatfield–McCoy feud involved two rural families of the West Virginia–Kentucky area along the Tug Fork of the Big Sandy River in the years 1863–1891. More than a dozen people were killed.

[https://en.wikipedia.org/wiki/Hatfield–McCoy\\_feud](https://en.wikipedia.org/wiki/Hatfield–McCoy_feud)