Fairness and Reciprocity

Armin Falk
University of Zurich, CESifo, CEPR, IZA

Berkeley, August 2002

Overview

- Why deal with fairness?
- Fairness and reciprocity in the lab
  - Bilateral games
  - Social dilemma games
  - Markets
- A field experiment
- Theories of fairness and reciprocity
- Evaluation of the theories
- How to proceed from here?

Why should economists take reciprocity into account?

- Reciprocity is real: Without a proper understanding of the nature of fair behavior and reciprocity our understanding of social reality is limited.
- Fairness is important for economic policy issues, e.g,
  - Labor compensation, wage rigidities
  - Optimal contract design, effectiveness of incentives
  - Social policy questions, legitimacy of the welfare state

Why has fairness and reciprocity largely been neglected?

- For a long time economists were preoccupied with perfectly competitive markets. In these markets fairness concerns are less important than in strategic interactions where agents can affect each others' payoff. Yet, many situations are not perfectly competitive.
- Game theoretic methods paved the way because they allow a precise analysis of strategic interactions and to model fairness concerns explicitly.
- Experimental methods paved the way for the empirical recognition of fairness motives.

Reciprocity…

- the reward of kind actions,
- the punishment of unkind actions,
- even if rewarding or punishing is costly.

Setting the stage: Moonlighting Game

(Abbink et al. 2000, Falk et al. 2000, Berg et al. 1997**)**

- 1. Stage:
  - Players receive an endowment of 12 points
  - Player A chooses action $a \in \{-6, -5, \ldots, 5, 6\}$
  - $a \geq 0$: A gives B a points
  - $a < 0$: A takes $|a|$ points from B
  - In case $a \geq 0$ the experimenter triplicates a such that B receives $3a$.
  - If $a < 0$ player A takes $|a|$ points from B and B loses $|a|$ points

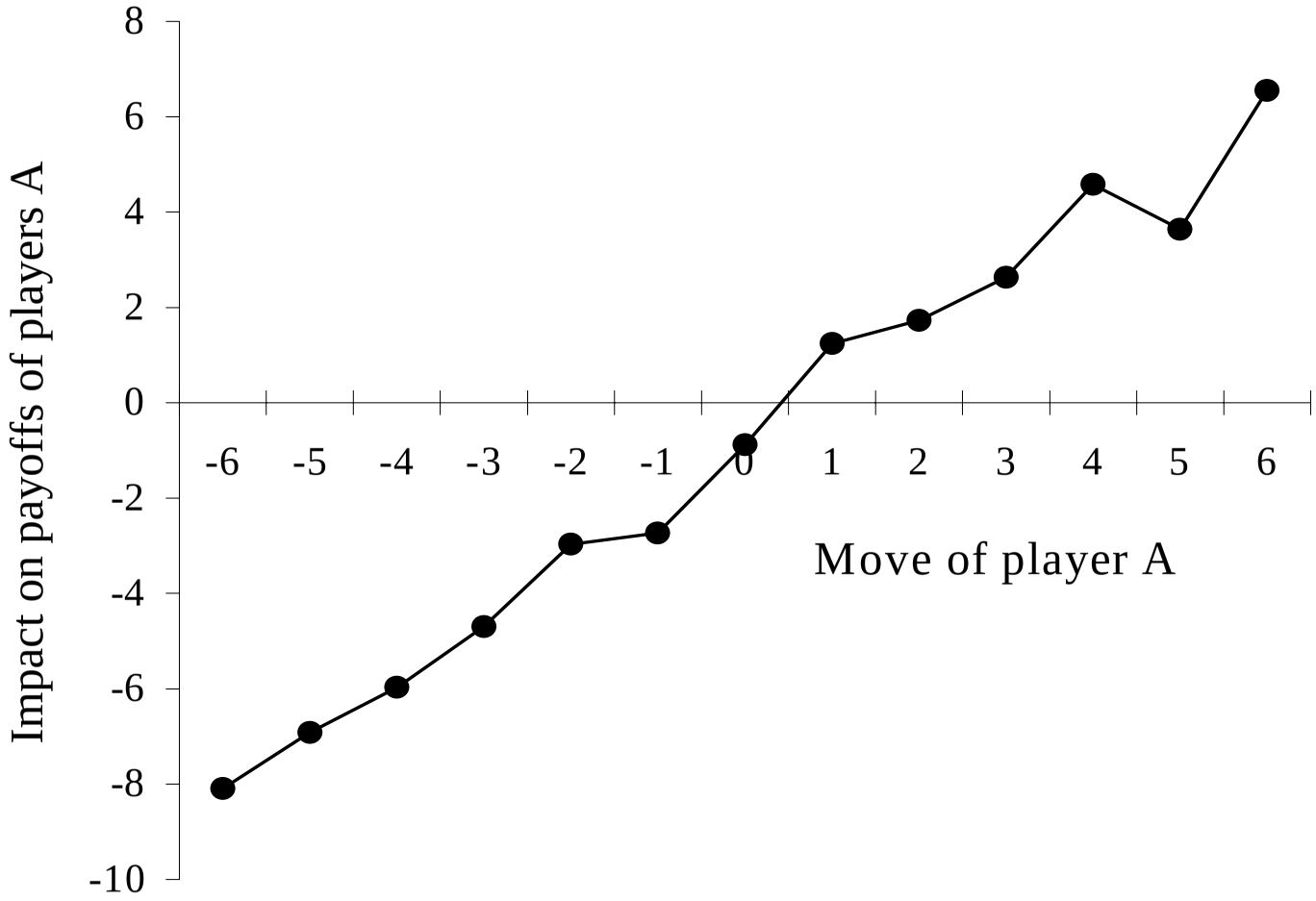
Moonlighting Game (ii)

- 2. Stage
  - B realizes a und chooses $b \in \{-6, -5, \ldots, 17, 18\}$
  - $b \geq 0$ is a reward for A
  - $b < 0$ is a punishment
  - A reward transfers b points to A
  - A punishment costs B $|b|$ points and reduces A‘s income at $3|b|$
- Prediction (selfish and rational): $b = 0$ for all a, and $a = -6$

Moonlighting Game (iii)

- Random allocation of roles
- Strategy method
- Anonymous one-shot interaction
- Experimental software z-Tree (Fischbacher 1999)
- 112 subjects (66 in the I-treatment and 46 in the NI-treatment), no economics students
- 1 point = 1 Swiss Franc (.65 US$).
- Subjects received on average CHF 22.20 in the I-treatment and CHF 24.10 in the NI-treatment (including a show-up fee of CHF 10).
- Experiment lasted approx. 45 minutes.

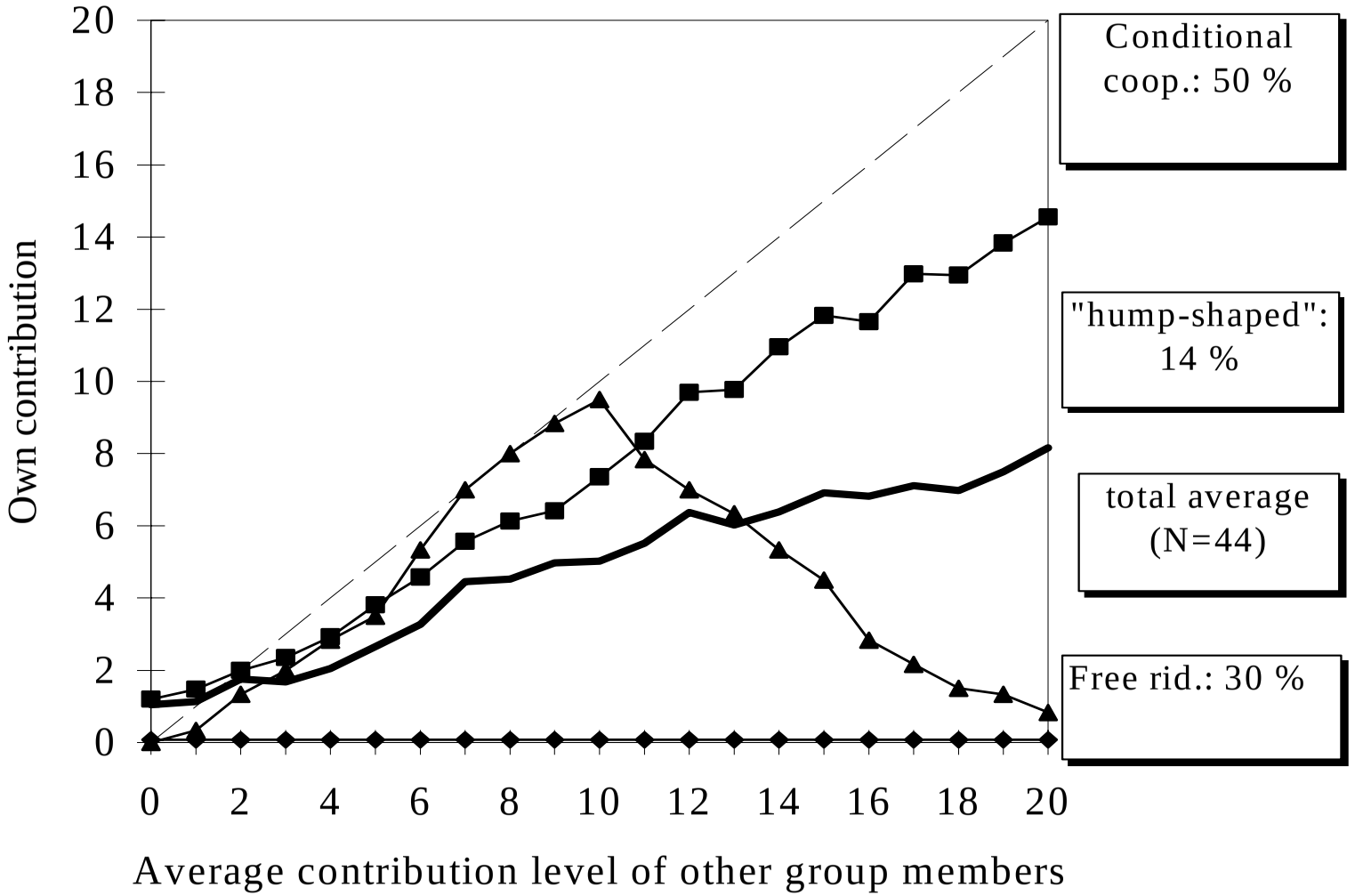
Moonlighting Game (iv)



Reciprocity in Social Dilemma Games

- How does fairness affect behavior in social dilemma games?
- Public Goods game where subjects can condition their contributions on the contributions of others (Fischbacher et al. 2001, Falk and Fischbacher 2002)
- 2 decisions, one unconditional one conditional, lottery
- Groups of four subjects. Each subject is endowed with $y = 20$ tokens. Subjects have to decide how many tokens to keep privately and how many tokens to invest in a group project.
- For each token invested in the project, **each** subject in the group receives 0.4 tokens, i.e., the group earns 1.6 tokens.
  $\Rightarrow$ Group as a whole benefits from a contribution.
  $\Rightarrow$ Yet, each contributor looses 0.6 tokens.
  $\Rightarrow$ Purely self-interested subjects will contribute nothing.

Experimental Results



Interaction of selfish and reciprocal players

- If selfish and reciprocal players interact, one would expect that eventually cooperation breaks down.
- Reciprocal players contribute conditional on what others do. Put differently: The only way to punish free riders is to withdraw contributions.
- Average contribution is between 40% and 60% during the initial periods.
- In the final periods about 75 percent of the subjects completely free-ride (meta study, F/S 1999).

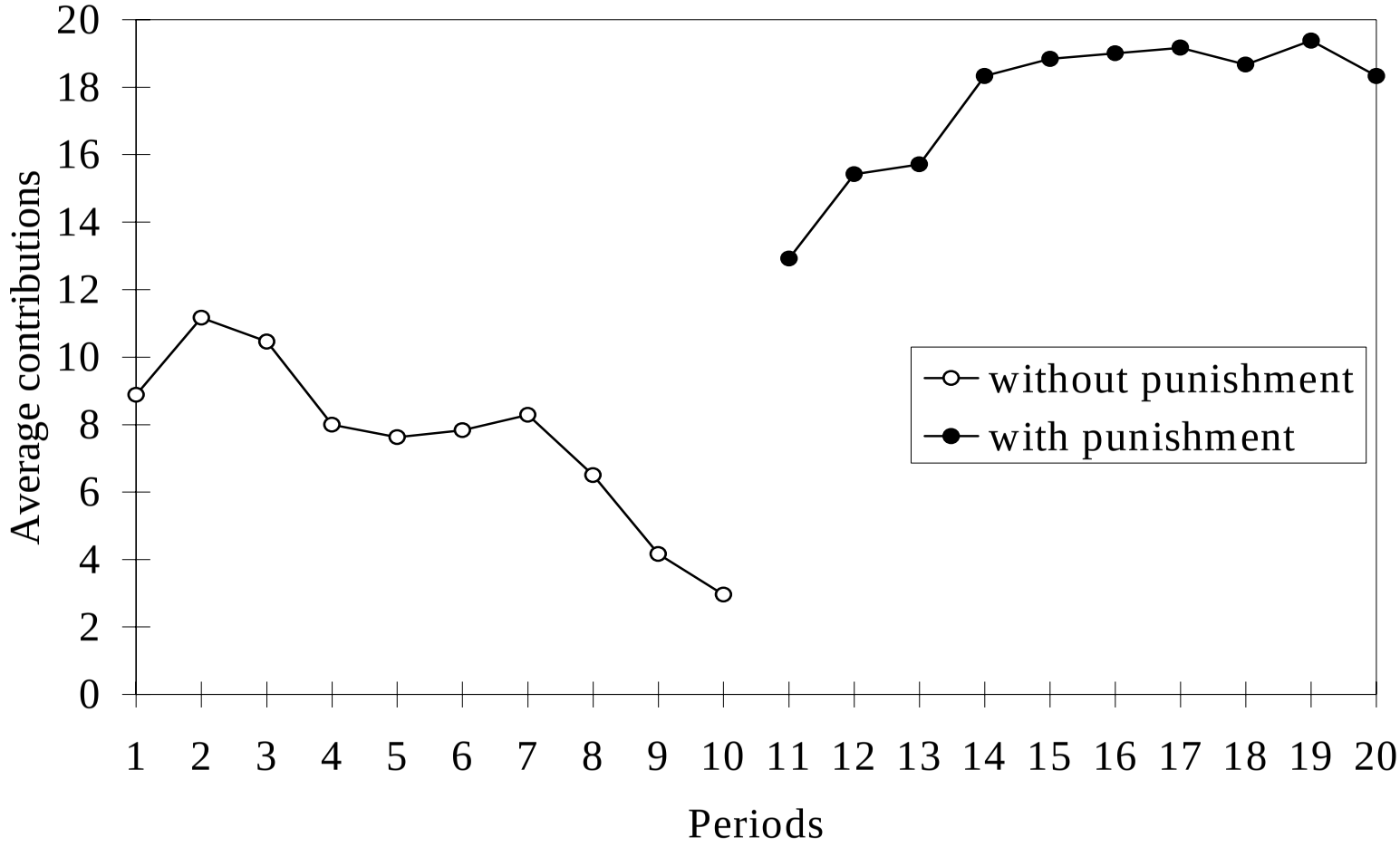
Interaction (ii)

- In a sparse environment, conditional cooperative players cannot achieve high contribution levels.
- What happens if they are given the chance to punish free-riders? (Fehr and Gächter 2000, Carpenter 2000, Falk et al. 2001)
- Stage 1: as above.
- Stage 2: Players decide simultaneously whether to assign punishment points to the other players after they observed (anonymously) how much the others contributed.
- Each punishment point reduces the Stage 1-Payoff of the punished subject by ten percent. Punishment is also costly for the punisher (roughly 1:3 relation)

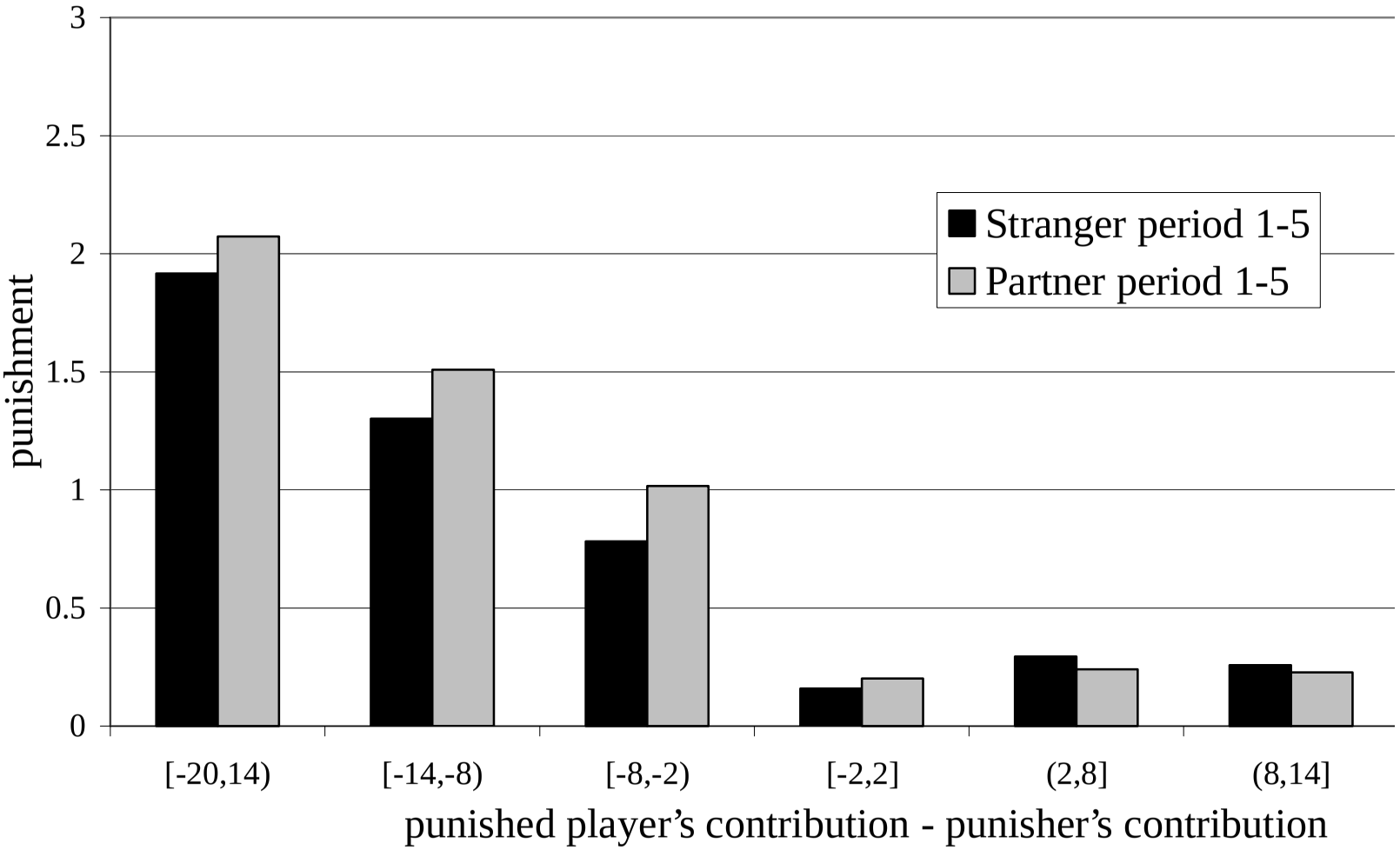
Interaction (iii)

- 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.

Experimental Results



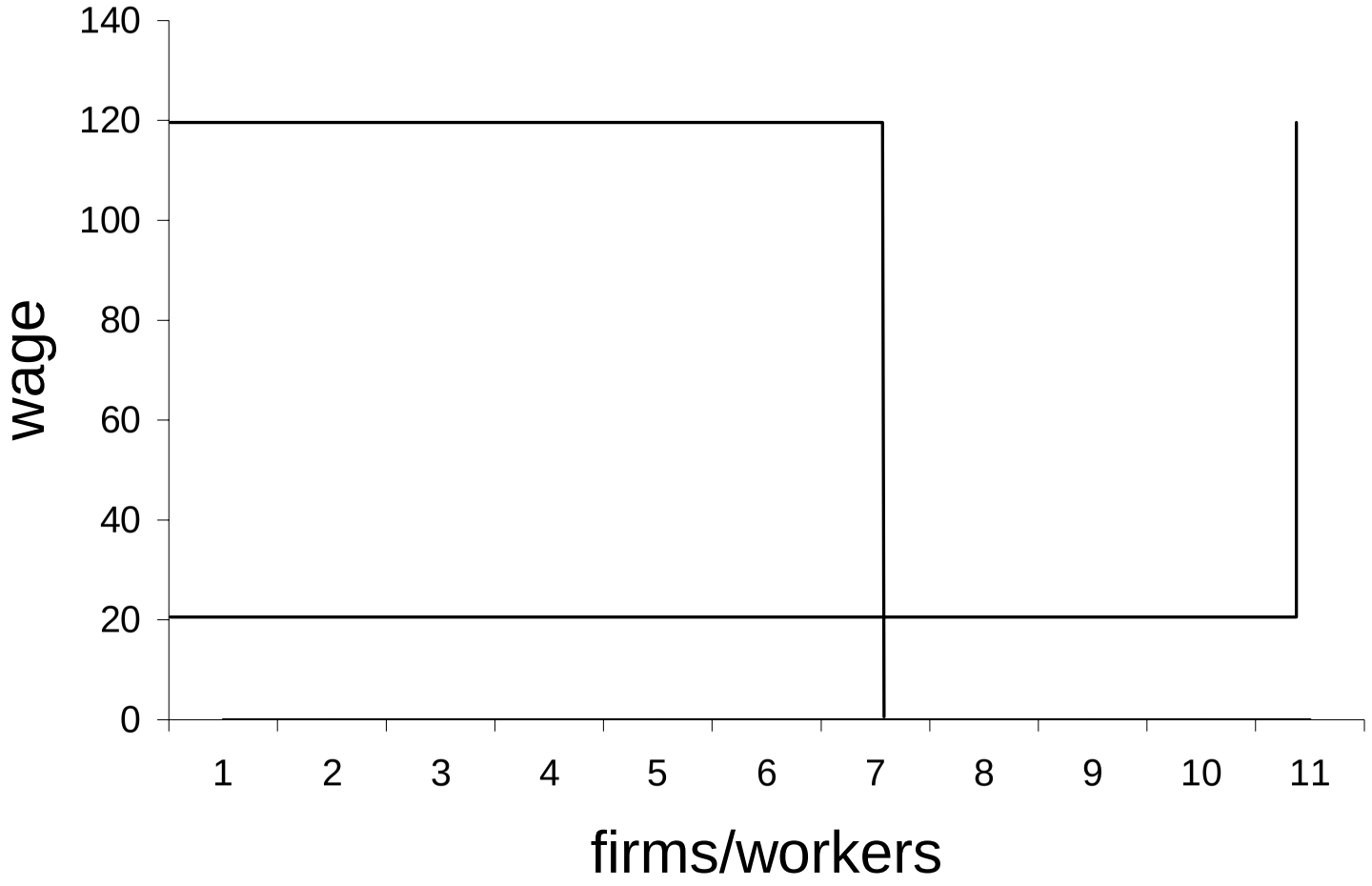
Punishment pattern in one-shot and repeated public goods gam (Source: Falk, Fehr, Fischbacher 2001)



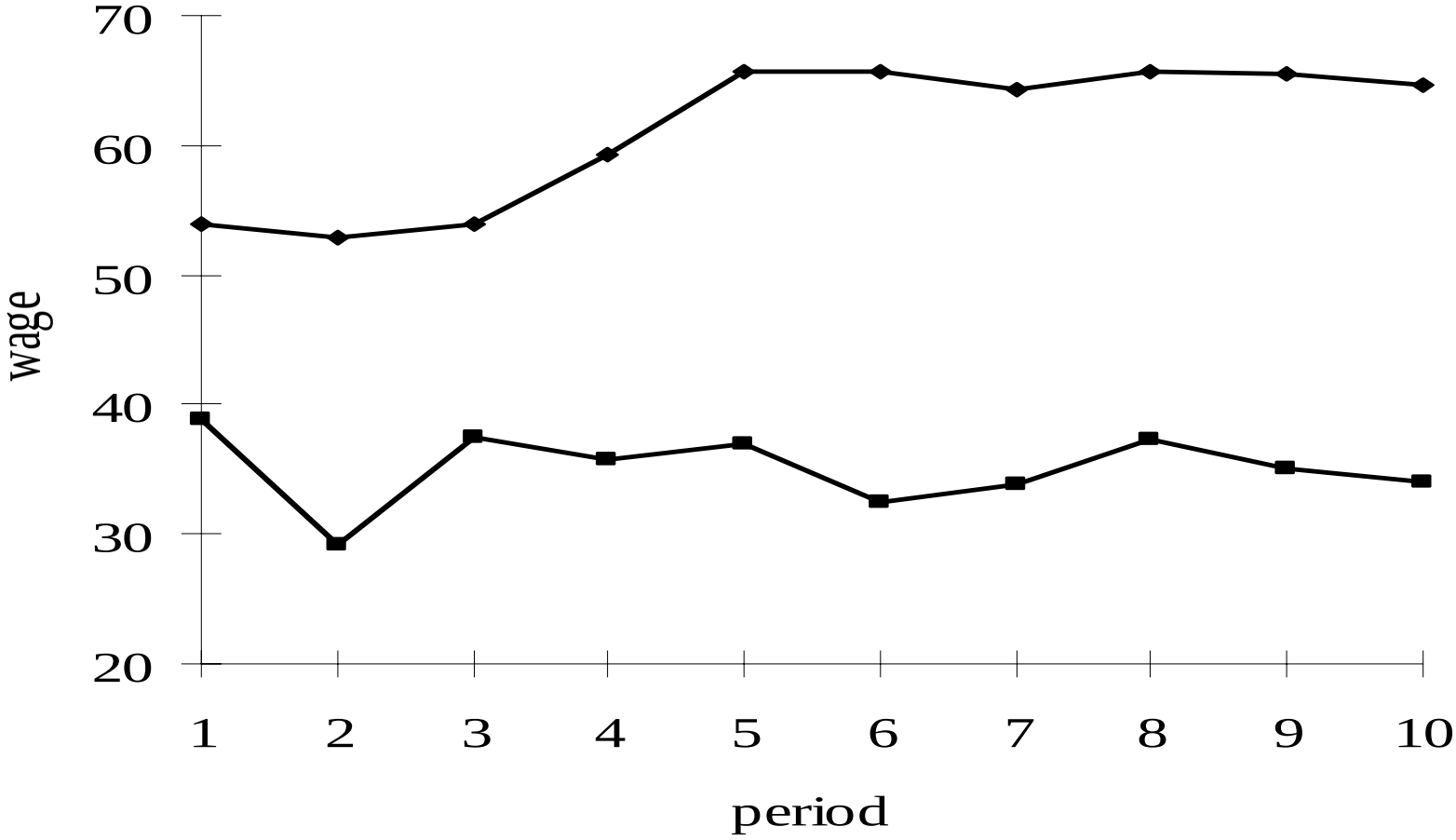
Reciprocity in markets

- Reciprocity is important in bilateral, multilateral and in market environments.
- The impact of reciprocity on the market outcome crucially depends on whether the market is complete or incomplete.
- Gift-exchange game (Fehr and Falk 1999)
- Stage 1: Firms and workers conclude contracts. Wages are settled in a double auction market, with wage $\in$ [20, 120]. There is an excess supply of workers (7:11). (UB = 20).
- Stage 2: Workers who concluded a contract choose an increasingly costly effort, with effort $\in$ [0.1, 1]
- Payoffs:
  - Firms: (120 – wage)effort
  - Workers: wage – cost of effort
- Standard prediction: wage = 20, effort = 0.1

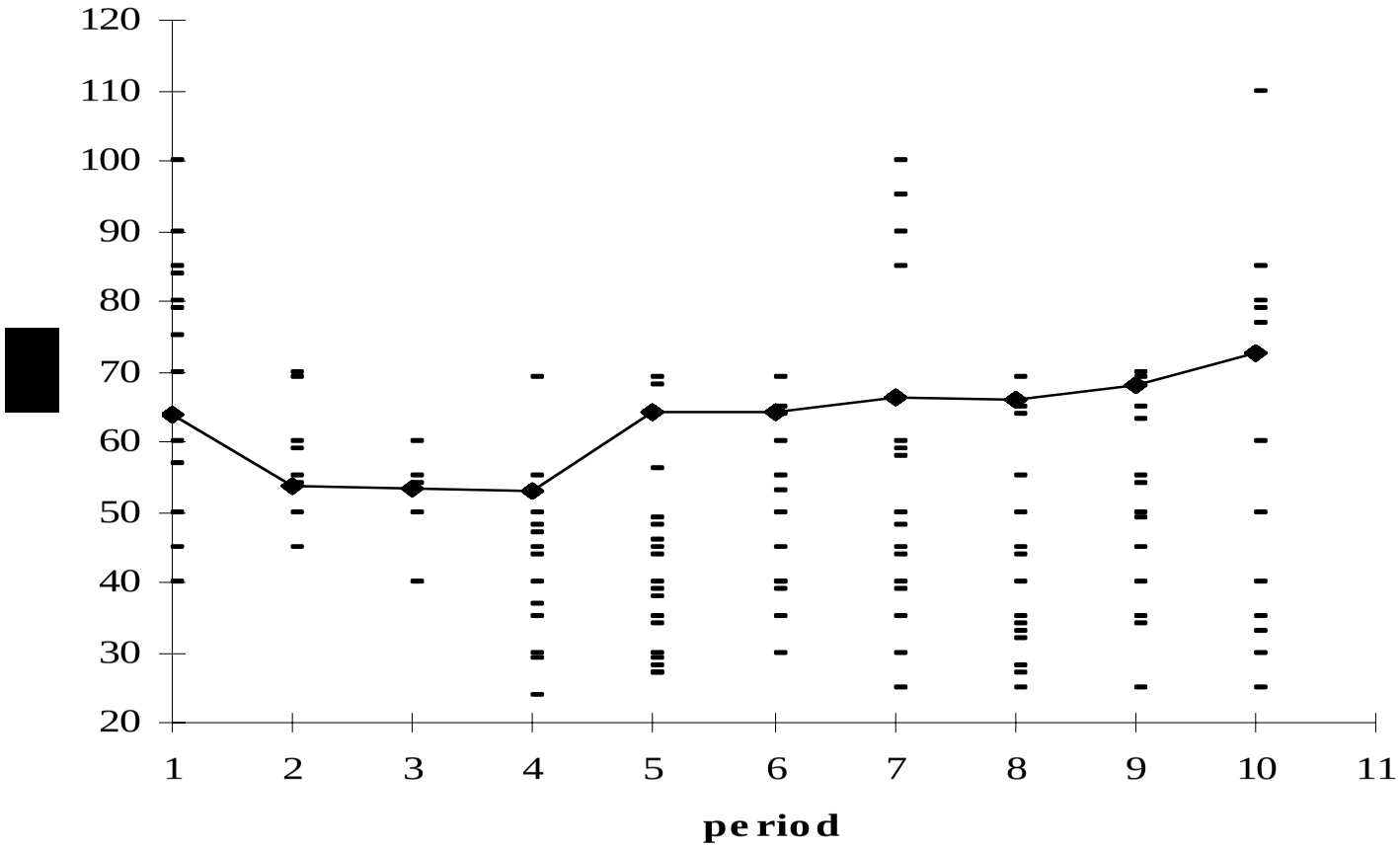
Competitive Prediction



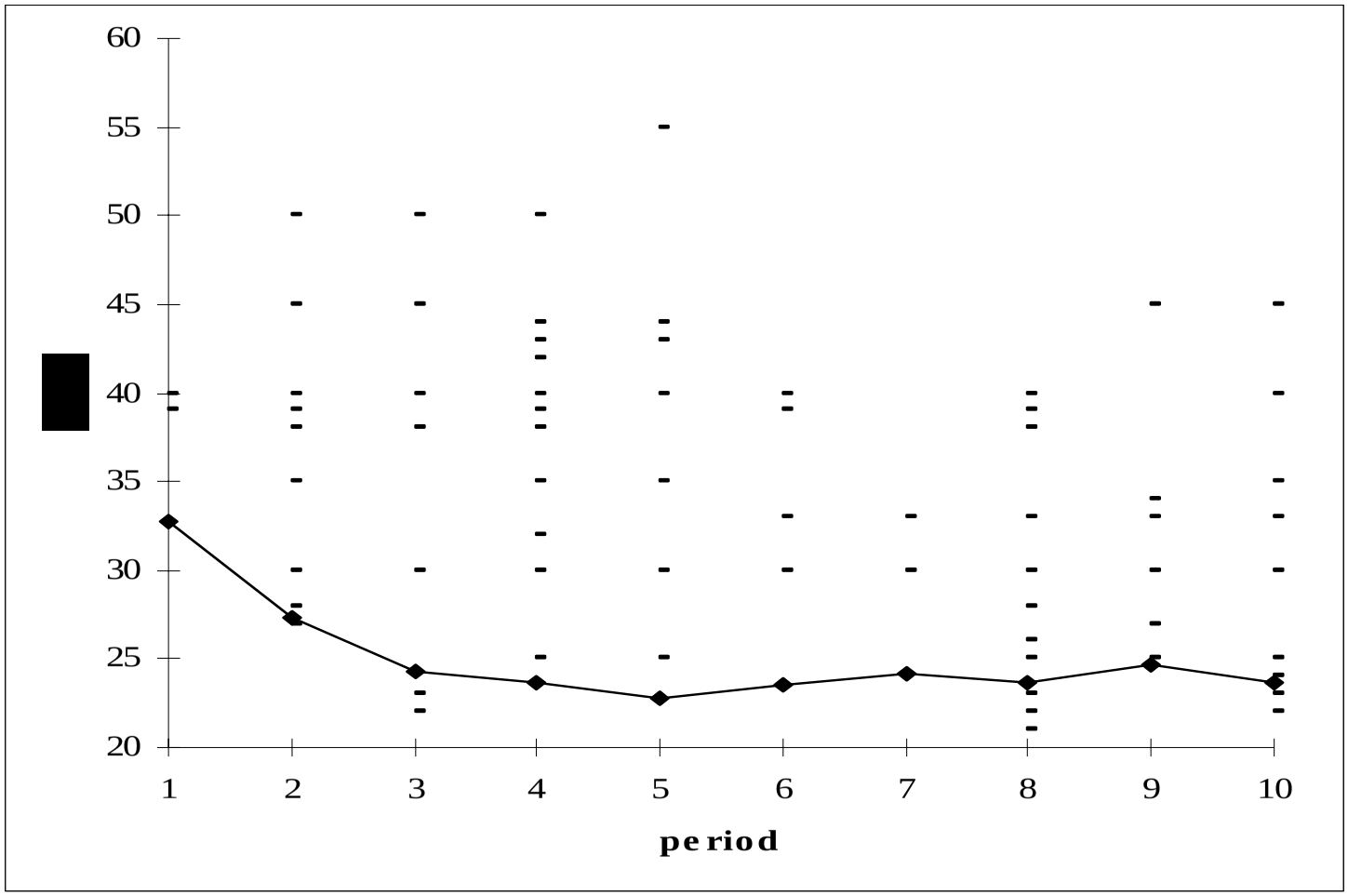
Reciprocity in Markets: Wages



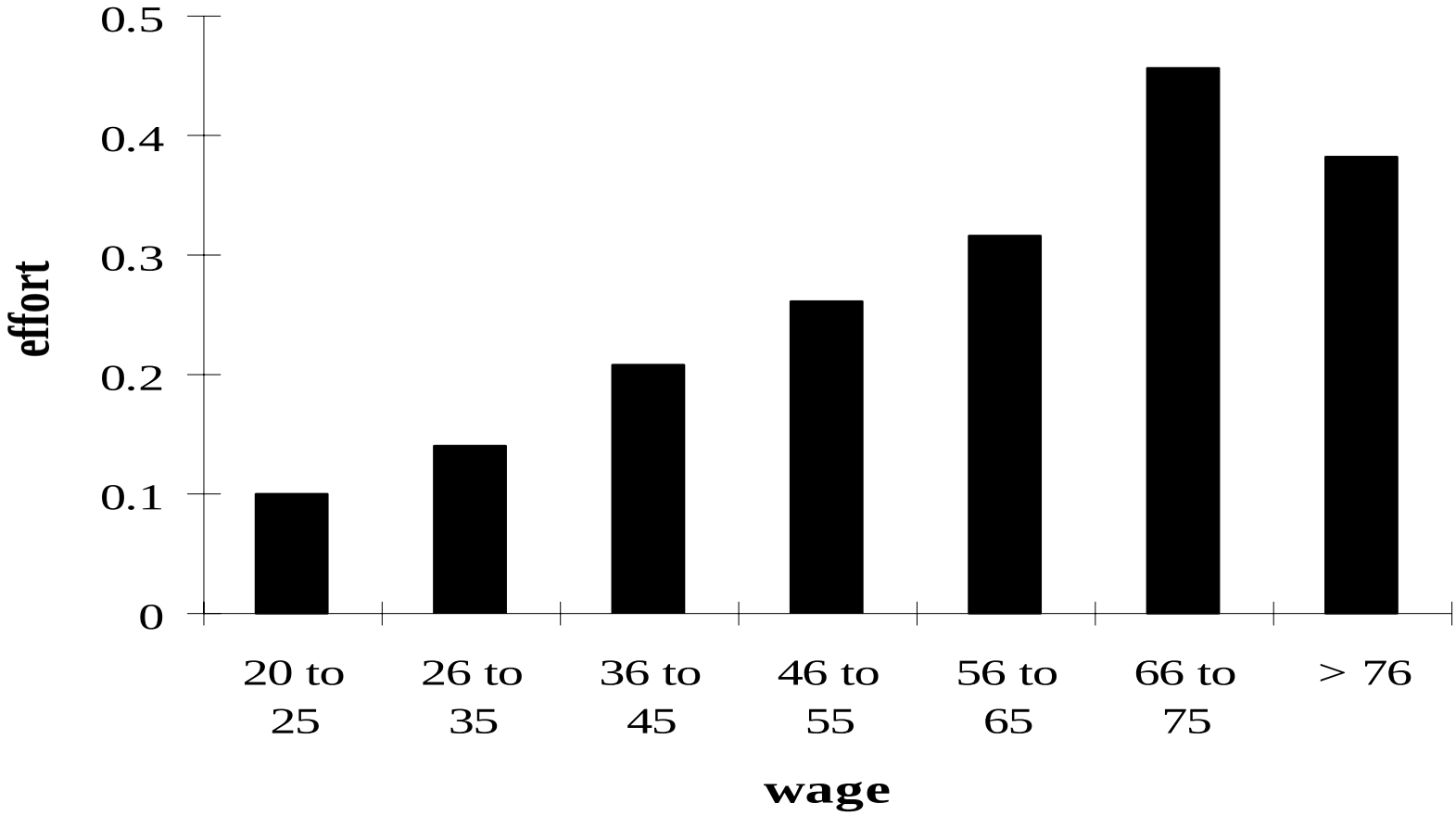
Underbidding: Incomplete Market



Underbidding: Complete Market



Reciprocity in Markets: Wage-effort Relation



Reciprocity in Markets

- In the incomplete contract market, wages are on average substantially higher than predicted.
- Underbidding of workers is not accepted by firms.
- Firms pay voluntarily high wages, because there is a positive correlation between wages and efforts on average.
- When effort is exogenously fixed, wages converge towards the predicted equilibrium and firms take advantage of underbidding.
- Reciprocity much stronger in repeated interaction (Gächter/Falk 2002)
- Reciprocity and endogenous long run relations in incomplete markets (Brown, Falk, Fehr 2002)

Questionnaire studies

- *“ In economics, it is normally assumed that people, being self-interested, must be either coerced or bribed into performing tasks. However, the main causes of downward wage rigidity have to do with employers’ belief that other motivators are useful as well, which are best thought of as having to do with generosity.”*(Bewley 1999, p. 431)
- Agell and Lundborg (1995) report that underbidding is not “all that uncommon”: 43 percent of firms had at least once encountered underbidding blue-collar and 53 percent underbidding white-collar workers.
- Firms refuse to employ underbidders: blue collar 95 percent and white collar 82 percent.
- ”While unemployed workers knock on the factory gates to a surprising extent, ... , firms keep them locked” (Agell und Lundborg, p. 299).

A field experiment

- An international charitable organization in Zurich that helps children in need around the world.
- Shortly before Christmas organization sends out letters with an appeal for charitable donations.
- The organization sends out letters to roughly 10.000 addresses.
- The money of the 2001 mailing was collected for homeless children in Vietnam (Dhaka, Bangladesh)
- Question: Does the willingness to donate depend on a gift included in the letter (“reciprocity”)?

Field study

- Three treatments:
  - No gift
  - Small gift (postcard painted by Vietnamese children)
  - Large gift (set of eight postcards painted by Vietnamese children)
- All addresses were randomly and evenly allocated to one of the three treatments.
- In the cover letter it was stated: “The postcards are a gift by the children of Dhaka in Bangladesh. You can keep it or give it to someone else.”
- Except for the gifts and these two sentences, everything was exactly the same across treatments.

Results

Donation across treatments

	no gift	small gift	large gift	no gift 00
number of letters	3262	3237	3347	9846
number of donations	397	465	691	1124
average number of donations	0.12	0.14	0.21	0.11
total of donations (< 500 CHF) in CHF	24,673	27,106	40,877	67,473

Treatment differences in the frequency of donations

Dependent variable: Frequency of donation

	Model 1	*Model 2*
Small gift dummy	0.022***	0.021***
	(0.008)	(0.008)
Large gift dummy	0.085***	0.081***
	(0.009)	(0.009)
Small gift x last year		0.047
		(0.036)
Large gift x last year		0.047
		(0.036)
Last year		0.243***
		(0.024)
Constant	0.122***	0.092***
	(0.006)	(0.005)
n	9846	9846
Prob. > F	0.0000	0.0000
R-squared	0.0098	0.0671

Note: The estimation procedure is an OLS-regression with robust standard errors (in parentheses).

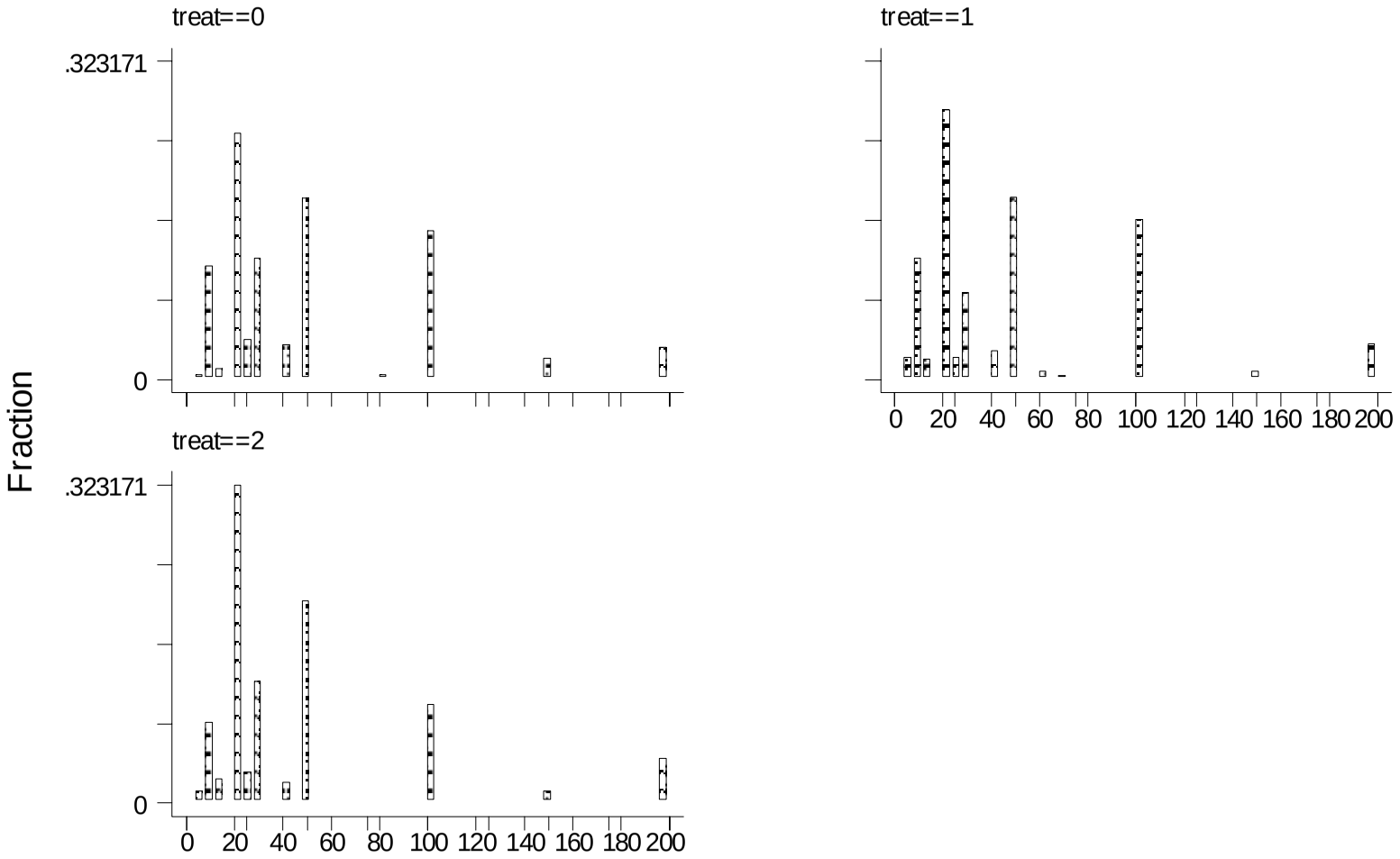
*** indicates significance on the 1-percent level.

Treatment differences in the absolute amount of donations

Dependent variable: Donations ≤ 500

Small gift dummy	13.391**
	(5.635)
Large gift dummy	46.200***
	(5.419)
Constant	-174.158***
	(5.811)
n	9807
Prob. > chi2	0.0000
Pseudo R-squared	0.0034
Log likelihood	11993.525

Note: The estimation procedure is Tobit. Standard errors are in parentheses. *** indicates significance on the 1-percent level, ** on the 5-percent level, respectively.



Histograms by treat

Some thoughts on gifts and charitable giving

- Total donations: 92,655 CHF
- Hypothetical total donations if all receive
  - No gift: 74,472 CHF
  - Large gift: 120,248 CHF
- Cost of gifts: ~2,000 CHF
- Actual net gain: 92,655 – 74,472 – 2,000 = 16,183 CHF
- Hyp. net gain: 120,248 - 74,472 – 6,000 = 39,776 CHF
- Did everybody like the gift...? Letters sent back ("neg. rec."):
  - No gift: 76
  - Small gift: 98
  - Large gift: 148
- Does it work next year in the same way?
- Does any gift do the job? (Here, gift is given from the receivers to the donators; symbolically creates a gift-exchange relation.)

Theories of Fairness and Reciprocity

- The experimental and field evidence is largely at odds with the standard economic assumption of narrow self-interest. As a response to the evidence, several theoretical models have been developed.
- All models assume that in addition to material self-interest, other regarding motives are crucial.
- **Altruism** (e.g., Becker 1974, Andreoni & Miller 1998): Other players' material payoff is positively valued. Cannot explain punishing behavior.
- **Relative Income and Envy** (Banerjee 1990, Bolton 1991): Other players' material payoffs are negatively valued. Cannot explain gift-giving and other nice behaviors.

Theories (ii)

- **Intention-based reciprocity** (Rabin 1993)
- Players reward kind and punish unkind intentions. Beliefs about other players actions, and beliefs about other players' beliefs about the own action, enter directly into the utility function; restricted to two person normal form games.
- Dufwenberg and Kirchsteiger (1999) extend and modify Rabin's notion of reciprocity to render it applicable to N-person extensive form games.

Theories (iii)

- **Inequity Aversion (**Fehr and Schmidt 1999, Bolton and Ockenfels 2000)
- Fehr-Schmidt: Other players' payoff is negatively valued if the others' are better off, and positively valued if the others' are worse off. This results from the aim of achieving equality.

$$_i = \pi_i - \frac{\alpha_i}{n-1} \sum_{j,\pi_j > \pi_i} (\pi_j - \pi_i) - \frac{\beta_i}{n-1} \sum_{j,\pi_i > \pi_j} (\pi_i - \pi_j)$$

- $\beta_I \leq \alpha_i$ (asymmetric inequity aversion)
- $\beta_I \leq 1$ (no money burning to achieve equality)
- Players are heterogeneous. A substantial fraction has FS-preferences, the rest is assumed to be selfish.
- BO: Utility is concave in material payoff and concave in *i*'s share of total income (max. at *1/N*)

Theories (iv)

- **Falk and Fischbacher 1999:** Combine inequity aversion and intention-based reciprocity (see also Charness and Rabin 2000).

$$U_i = x_i + \rho_i \Sigma \varphi_{ji}(n)\sigma_{ij}(n)$$

- ρ_i is an individual reciprocity parameter
- The kindness term φ_{ji} measures how kind player *i* is treated by player *j* at node *n* – in the view of player *i*.
- The reciprocation term $\sigma_{ij}(n)$ measures the impact of player *i*'s reciprocal action, i.e., how *i*'s action alters *j*'s payoff (basically player *j*'s payoff).
- If the kindness term φ_{ji} is positive, player *i* increases his utility by increasing *j*'s payoff – if the kindness term is negative, player *i* increases his utility by decreasing *j*'s payoff.
- The product $\varphi_{ji}(n)\sigma_{ij}(n)$ is summed up over all nodes at which player *i* has to make a choice.

Theories (v)

- **How kind is - from *i*'s perspective - player *j*?**
- A strategy of *j* is perceived to be kind by *i* if it causes a payoff for *i* which is higher than the payoff of *j*. This is different from Rabin and Dufwenberg and Kirchsteiger who define "kindness" in relation to the feasible payoffs of player *i* and not in relation to the payoff of player *j*.
- Is the favorable distribution for *i* caused intentionally or not? If so, $\varphi_{ji}(n)$ is larger. However, even if player *j* is a dummy player who has no choice to make, the kindness term $\varphi_{ji}(n)$ is positive. It then reflects pure inequity aversion.
- Player *j*'s action is caused intentionally if there are 'reasonable' alternatives *j* could have chosen.
- Thus, kindness is determined by (i) the outcome caused by an action and (ii) the action's underlying intentions.

Evaluation of different theories

- Are sanctions driven by inequity aversion or the desire to retaliate (reciprocity) ?
- How important are intentions?

Sanctions: Reciprocity vs. Inequity aversion

- Three person one-shot public goods game with punishment opportunity:
- 1st Stage: public goods game
  - Contribute 20 points (cooperate) or nothing (defect)
  - Payoff
    - 20 - own contribution +
    - 0.6 * sum of all contributions
- 2nd stage: Reduce the other player's payoff at a cost
- 1 point reduction costs 1 points, i.e., inequity cannot be reduced by punishing

Payoffs on 1. Stage

	Both others defect	One of the others defects	Both others cooperate
Player i defetcs	20	32	44
Player i cooperates	12	24	36

Predictions

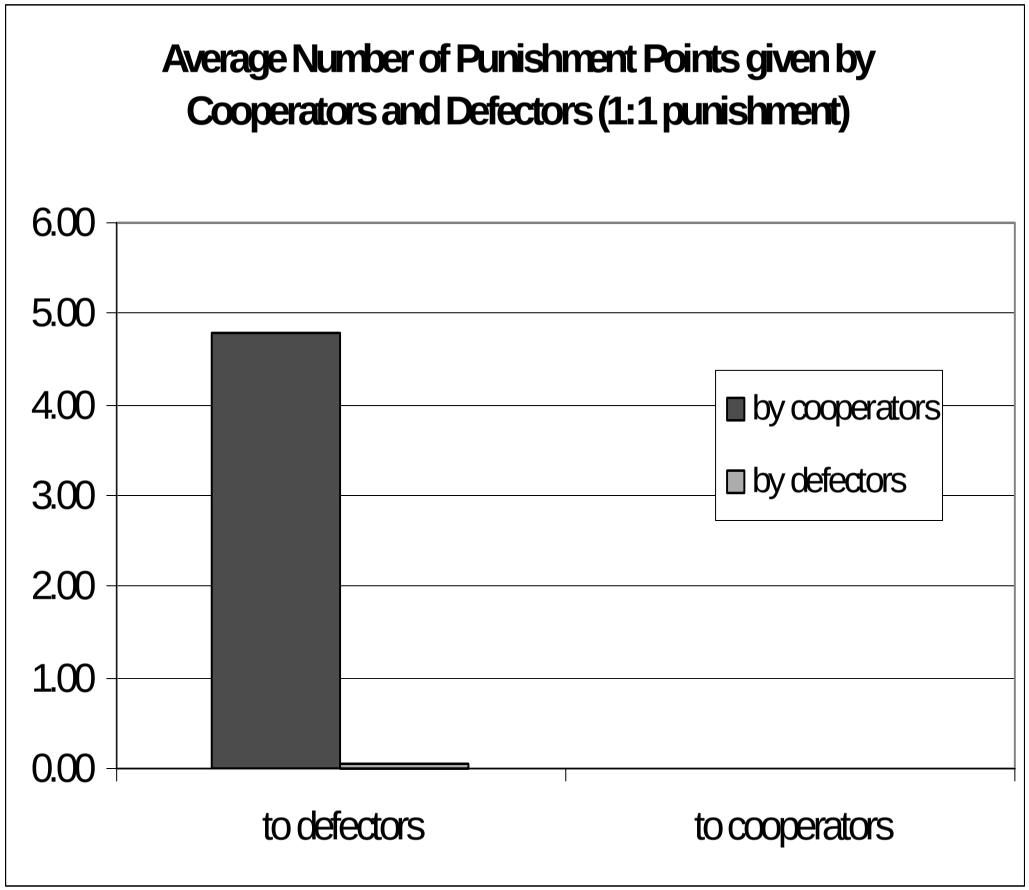
- If everybody is a selfish money maximizer:
  - No sanctions because they are costly.
  - No cooperation because defection is a dominant strategy.
- Fairness: Someone who defects acts in an unfair manner. Those who cooperate will therefore sanction defectors.

Are sanctions driven by inequity aversion?

- Inequity aversion:
  - People dislike payoff inequity and reduce it, even at a cost (Fehr/Schmidt 1999, Bolton and Ockenfels 2000).
  - Inequity aversion models predict no punishment because inequity cannot be reduced
- Reciprocity:
  - People reward fair and sanction unfair behavior, even if this is costly (Rabin 1993, Falk and Fischbacher 1999, Dufwenberg and Kirchsteiger 2000).
  - Reciprocity models predict sanctions because regardless of the cost of punishment and the possibility to reduce payoff inequity, defection is an unfair treatment, which „deserves" punishment.

Inequity aversion vs. reciprocity

N =120

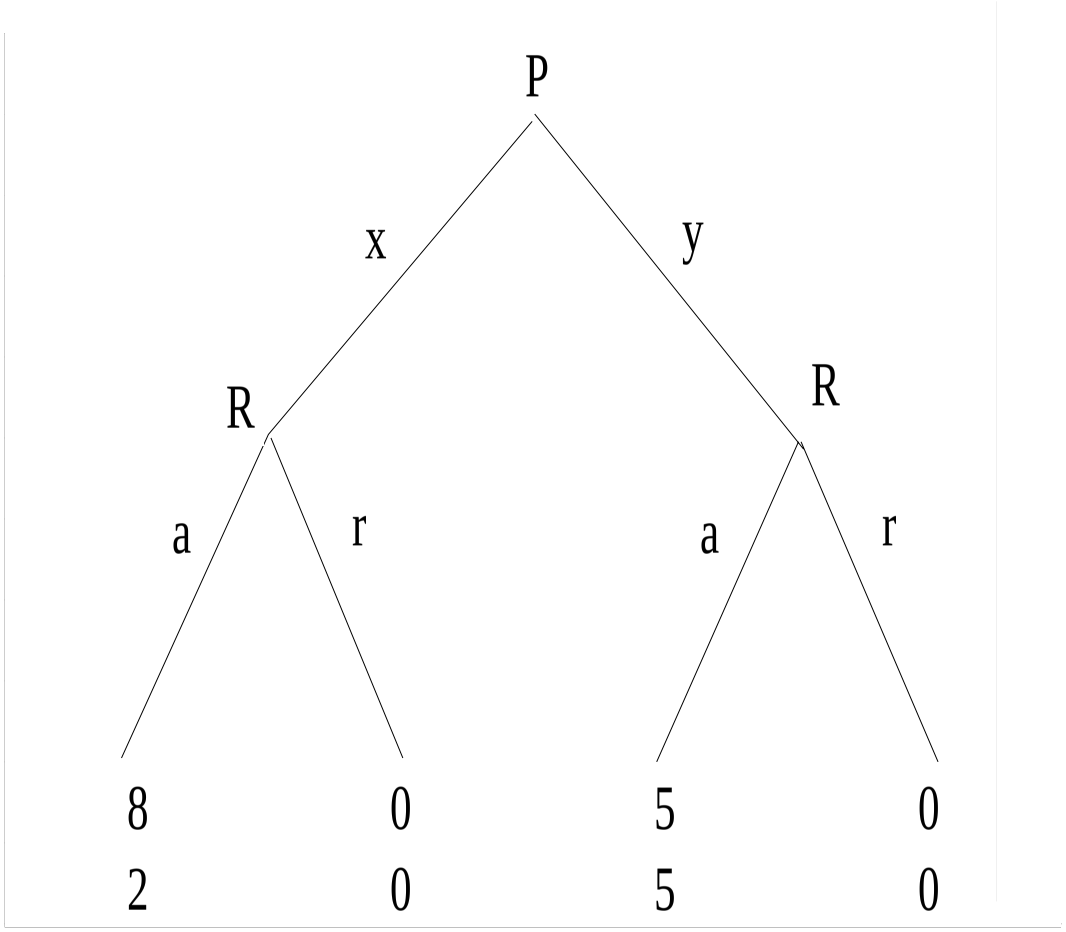


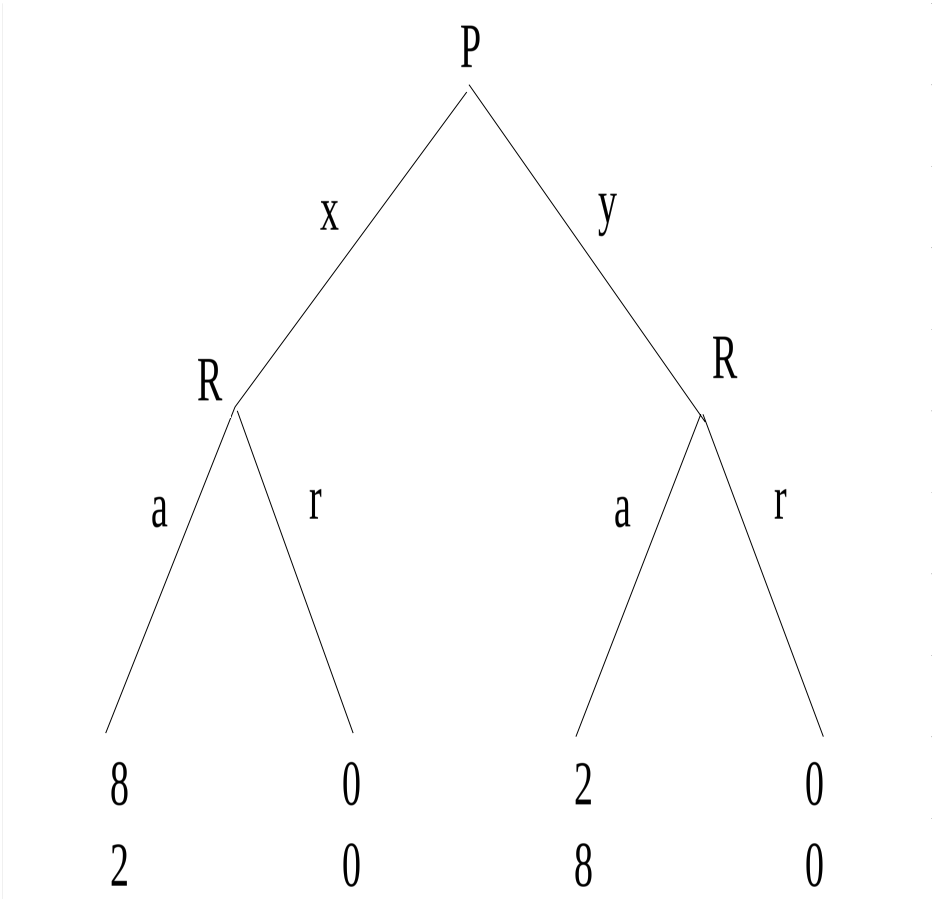
- 51 percent cooperate
- 47 percent of these cooperators punish two defectors.
- Punishment behavior of about 24 percent is incompatible with any inequity aversion model
- Sanctions and cooperation vary with cost of sanction.
- However, in high cost treatment, subjects spend more money on sanctions.

The role of intentions

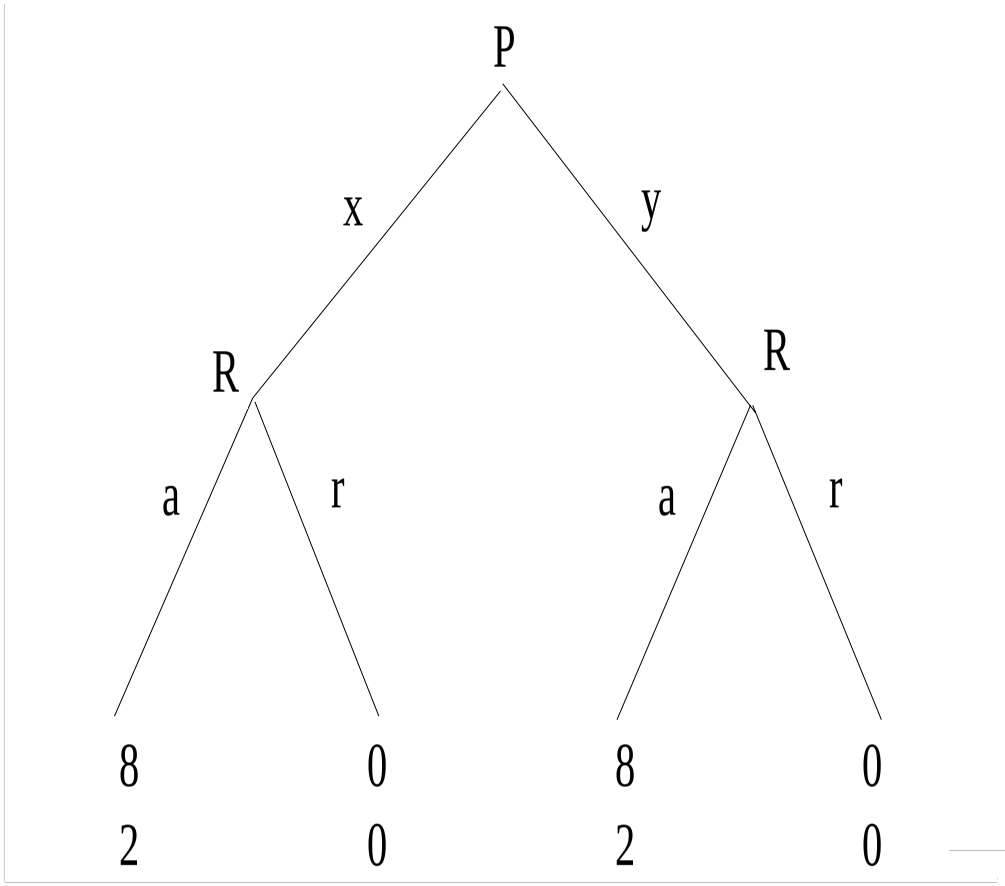
- The signaling of fair or unfair intentions rests on two premises:
  - The strategy space allows for fair and unfair actions.
  - The action is under the full control of the person who performs it.
- If intentions are behaviorally relevant, sanctions should be the stronger, the more unfair the intentions are.

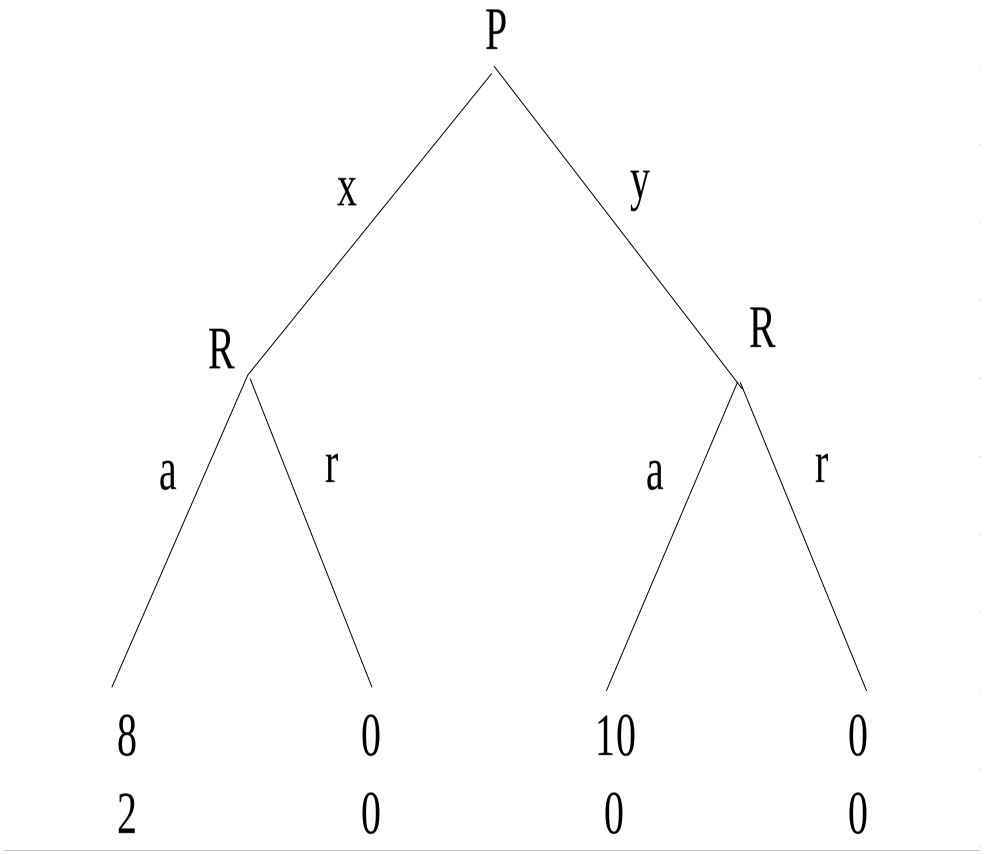
The role of intentions: Four Mini Ultimatum Games





Intentions (ii)

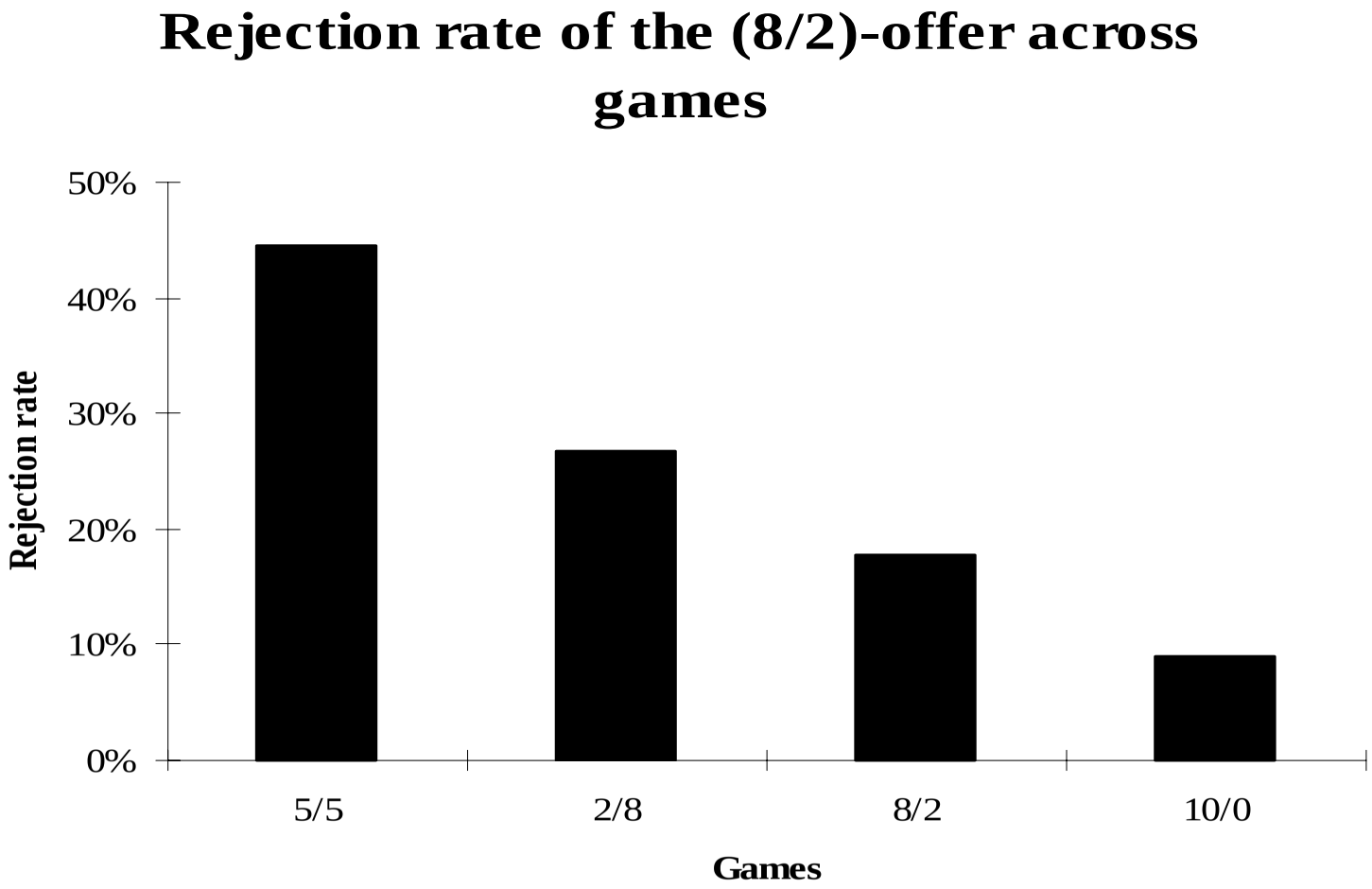




Predictions of the rejection rates of the 8:2 offer

- Inequity aversion models of BO and FS predict the same rejection rate for all games.
  - These theories model fairness in a consequentialistic way and the consequence of the 8:2 offer is always the same.
- Reciprocity models of R, DK model the fairness of an action as dependent on intentions; FF model: intentions and outcome
  - Different rejection rates are predicted.

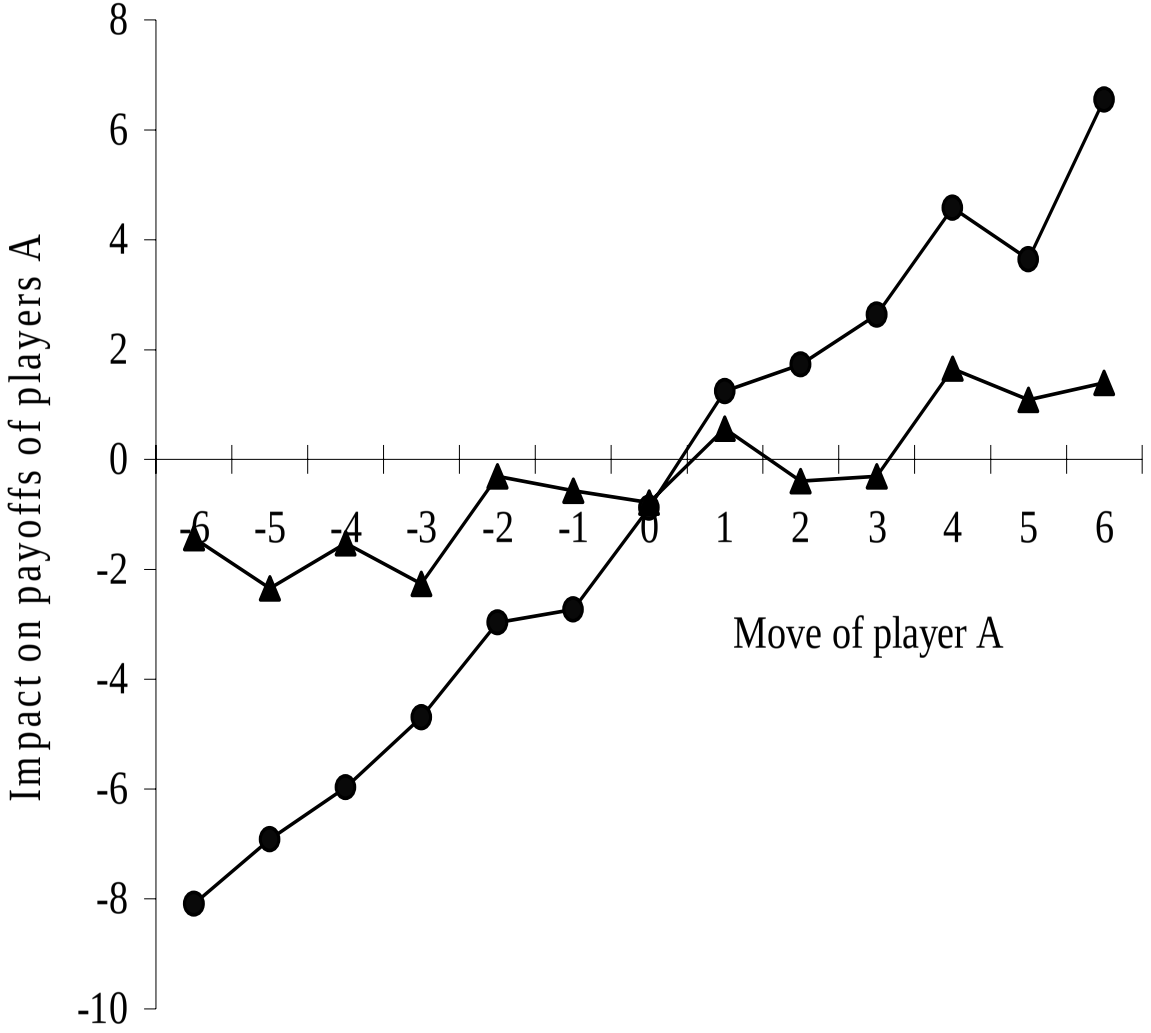
Experimental results (N=45)



Intentions, once more

- Moonlighting game as before.
- But: Player A‘s decision is randomly determined and players B know that.
- Random mechanisms is based on a „human choice distribution“. Controls for the equality of choice probabilities across computer generated and and human generated first-mover action.

Rewards and punishments with and without intentions



- The same consequences trigger very different behavior.
- Questions consequentialistic notions of fairness.
- Casts doubt on the consequentialistic practice in economics to define the utility of an action solely in terms of the consequences.

Which model?

- Most models predict many experiments rather well.
- The predictive power of the inequity aversion is limited (intentions, punishments if equity cannot be reached). However, they are quite simple.
- Pure intention-based theories have in general multiple equilibria even in the simplest games. They underestimate the importance of outcomes.
- Models that combine inequity and intentions capture the evidence best.

How to proceed from here?

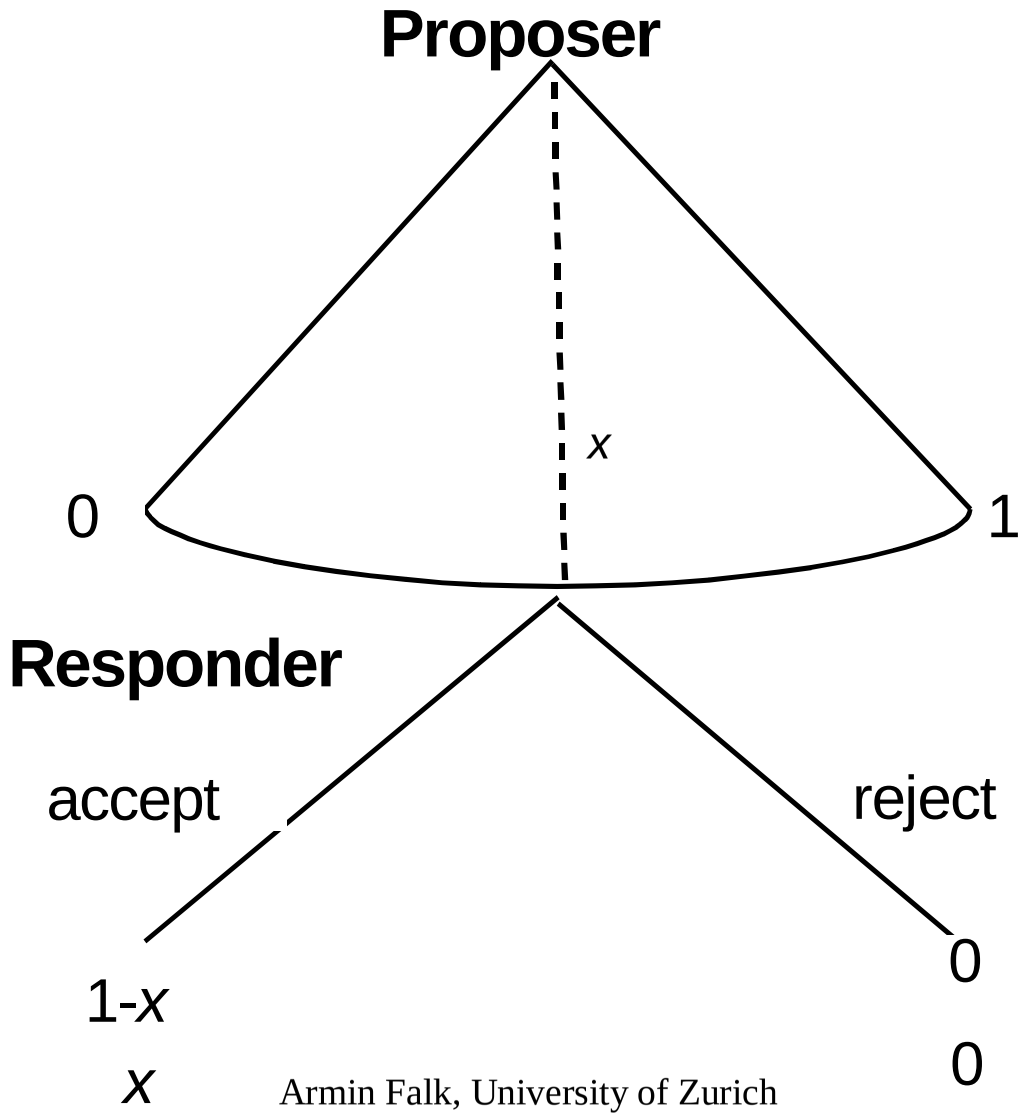
- Better models, which are simple and get the right predictions?
- More experiments are required to better understand the interaction between reciprocity and institutional environments (horizontal and vertical, production and distribution, markets).
- The development of alternative economic policy advice. Today's policy advice is almost exclusively built on the assumption that all people are selfish. Given the presence of fair and selfish types this advice is likely to be misleading.

Some examples

- Tax compliance: Conditional cooperation with other tax payers, the state, the tax authorities
- In the presence of conditional cooperation, there are multiple equilibria: belief-management (suppressing public disorder, advertisements).
- Reciprocity as a source of informal sanctions: key to the enforcement of implicit agreements and social norms. Part of a society's social capital. Should be supported and used by policy (danger of undermining).
- Labor compensation and wage rigidities.
- Optimal contract design, effectiveness of incentives.
- Social policy questions, legitimacy of the welfare state.

Additional Material

Ultimatum Game (Güth et al., 1982)



Ultimatum Game (ii)

- Prediction
  - Responder accepts $x \geq 0$
  - Proposer offers $x = 0$, which is accepted
- Facts
  - Virtually no offer above $x = 0.5$
  - Vast majority of offers between 0.4 and 0.5
  - Responders frequently reject offers $x < 0.2$
  - Facts hold across culture and stake size

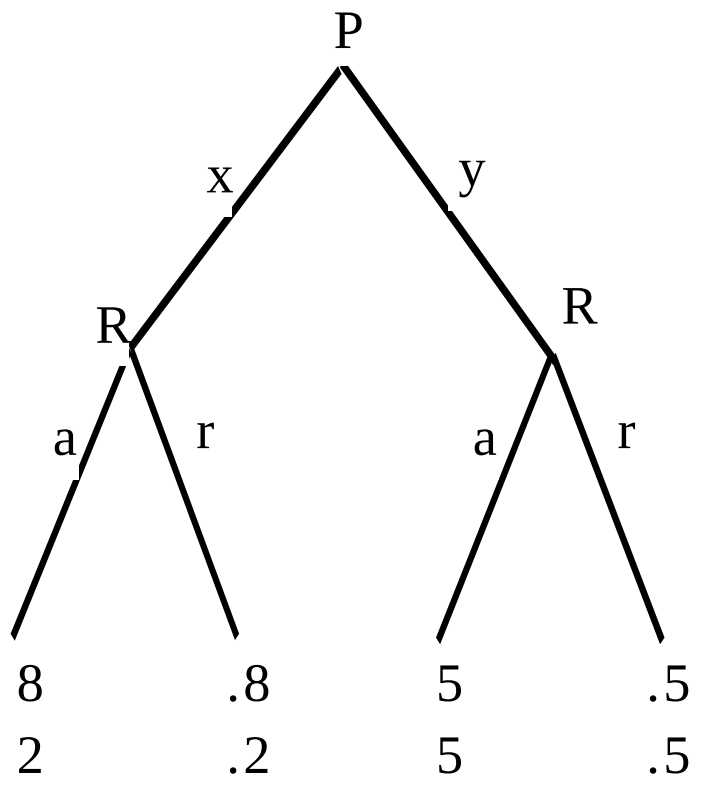
Ultimatum Game and Reciprocity

- Most responders do not act according to the *homo oeconomicus* assumption. They prefer to reject positive amounts of money rather than to accept an unfair treatment.
- **Reciprocity**: The reward of kind and the punishment of unkind actions, even if rewarding or punishing behavior is costly.

Or is it altruism?

- Why do proposers offer „fair“ offers?
- Altruism or fear of rejection?
- **Dictator game**: Responder must accept every offer (e.g., Forsythe et al. 1994)
- Results: Proposers make significantly lower offers.
- There is not much unconditional altruism.

Is punishment driven by inequity aversion?

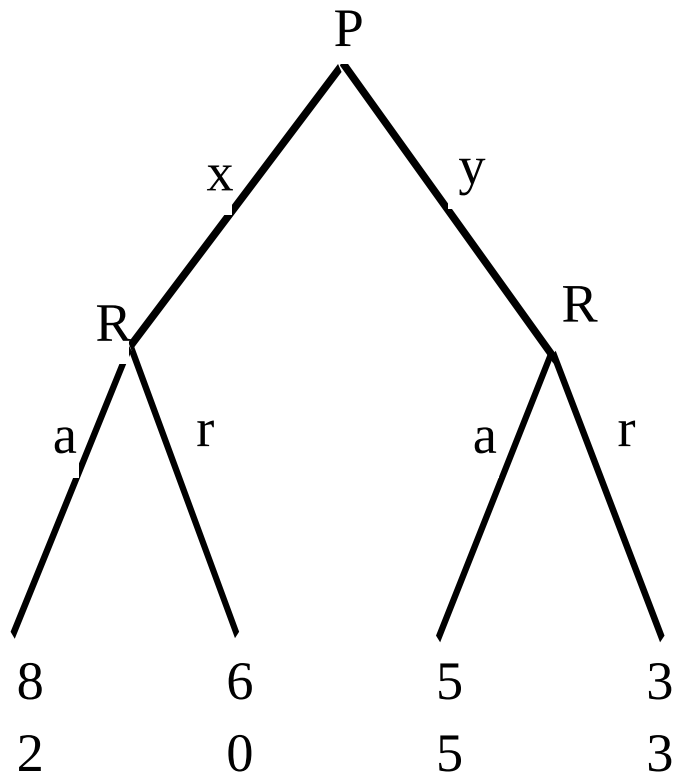


UG with constant relative share

- Rejection reduces payoffs to 10 percent
- Rejection cannot change the relative share

- Hence, BO predict no punishment
- The other theories predict rejections

Is punishment driven by inequity aversion?



UG with constant difference

- Rejection reduces payoffs by 2 points
- Rejection cannot change payoff differences

- Hence, FS and BO predict no punishment
- DK and FF predict rejections
  - 8:2 is unkind and triggers punishment. Punishing means a reduction of the other player's payoff.

Experimental results (N=48)

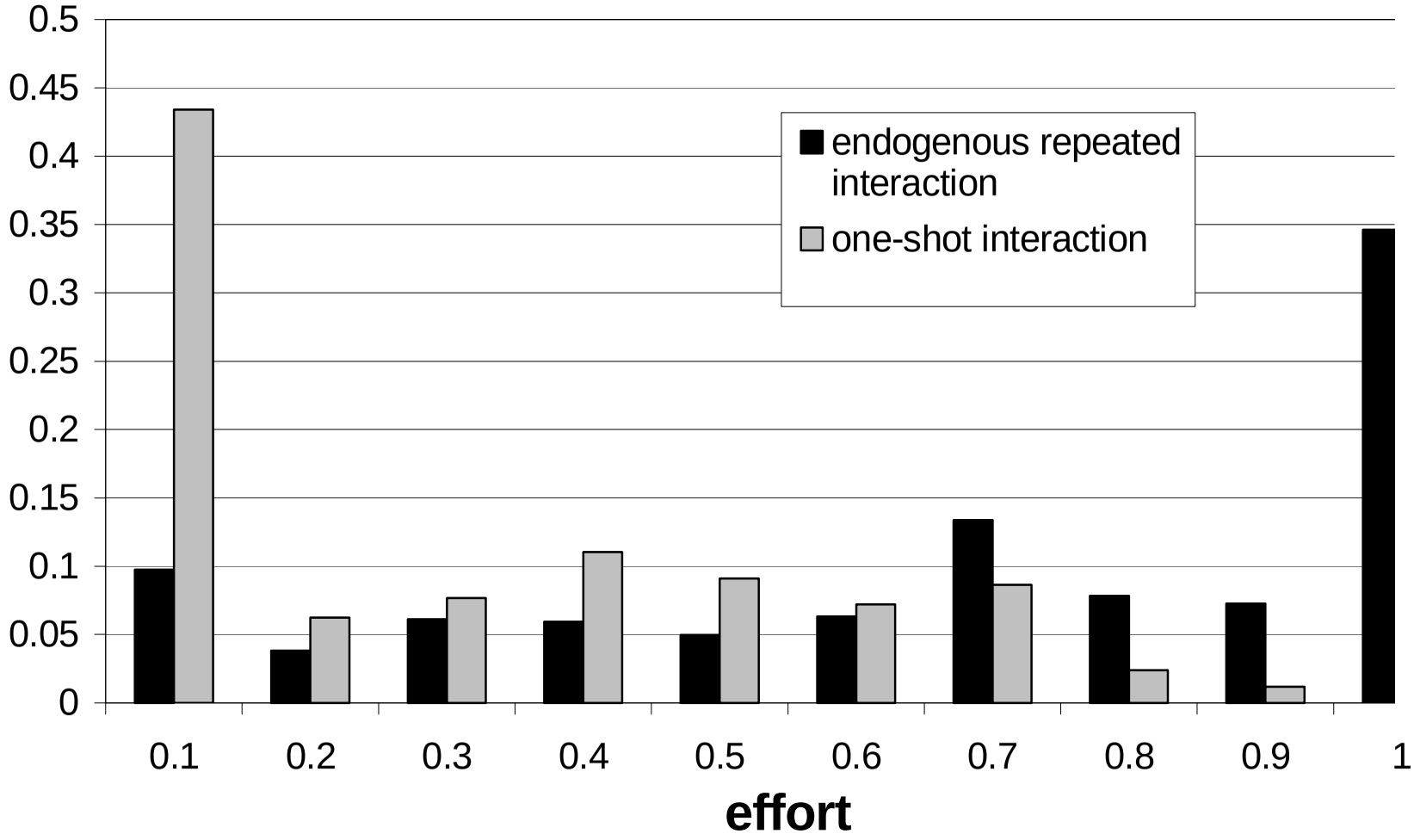
	Predict rejections of 8:2 offer				Result Rejection rate
	BO	FS	DK	FF	
UG with constant relative share	no	yes	yes	yes	38%
UG with constant difference	no	no	yes	yes	19%

- Punishment does not only occur to reduce inequity. Even if inequity cannot be reduced, people punish to reciprocate unkindness (20 percent).

Reciprocity and social capital

- Informal sanctions: Sanctions that are non enforceable by third parties (e.g., law/courts) and are therefore not part of a formal and enforceable contract or agreement.
- They are key to the enforcement of implicit agreements and social norms.
- The importance of informal sanctions derives from the fact that the bulk of people's daily interactions is not governed by explicit, enforceable contracts but by informal agreements and social norms.

Distribution of effort in one-shot and endogenously repeat gift exchange games (Source: Brown, Falk and Fehr)



Who is the relevant reference agent?

- Three person one-shot public goods game with punishment opportunity as above:
- 1st Stage: public goods game
  - Contribute 20 points (cooperate) or nothing (defect)
  - Payoff
    - 20 - own contribution +
    - 0.6 * sum of all contributions
- 2nd stage: Reduce the other player's payoff at a cost
  - Punishing cooperators: 1 point reduction costs .3 points.
  - Punishing defectors: 1 point reduction costs .4 points.
  - It is cheaper to punish cooperators.

Predictions

- BO models the fairness relation between an individual and the group average (fair share).
- In a situation where a cooperator faces a cooperator and a defector they predict that if cooperators sanctions he sanctions the other cooperator. It is the cheapest way to reduce inequity.
- The other theories model the fairness relation as a comparison between own payoff and the payoff of each other player.
- They predict that if cooperators punish, they punish defectors. Either because they have a higher payoff (FS) or because they are unkind (DK and FF).

Experimental Results

Allocated deduction points of cooperators