Fairness

- Experimental evidence for fairness
- Models of fairness
- Testing fairness models

Fair Behavior

- **Preferences for fair (equal) outcomes**
    - Not conditional on behavior
    - Relative to a reference standard
- **Reciprocity**
    - Positive reciprocity
        - Rewarding kind behavior
    - Negative reciprocity
        - Punish unkind behavior

- **Important Implication:**
    - ➢ Intervention to adjust outcomes, or punish unkind behavior, can be rational, even if it is costly.

Evidence for Fairness

- **Field evidence**
  - Collective action (strikes, consumer protest, voting)
  - Tax compliance (people pay more than is optimal given they are rational and selfish)
  - Donations to charity

- **Questionnaire studies in labor market**
  - Bewley (1995, 1997)
  - Agell and Lundborg (1995)
  - Campbell and Kamlani (1997)

Experimental Evidence for Fairness

- **Bargaining**
  - Equal offers in bargaining games
  - Disadvantageous counter offers
  - Ultimatum game: rejections of positive offers
  - Dictator game: positive transfers
- **Trust game, moonlighting game, and gift exchange game (this last game is in a market setting)**
  - Positive and negative reciprocity
- **Public Goods Games**
  - Cooperation higher than predicted by standard theory
  - Conditional cooperation
- **Punishment (in public goods games)**
  - Punishment of defectors

Trust game / Investment game

Berg, Dickhaut, McCabe (1995)

- 2 player sequential game, P1 and P2.
- Both players are endowed with 10 points
- P1 can give P2 up to 6 points (investment).
- Each invested point is tripled.
- P2 is informed about the investment, and can give points back to P1 (but does not have to)
- Standard prediction:
    - P2 gives nothing back to P1 (independent of investment)
    - Therefore P1 invests nothing

Results

Berg, Dickhaut, McCabe (GEB, 1995)

- P1's do invest

- P2's give back points
- P1 benefits from investments of 5 and 10
- On average P1's are just compensated

  - Evidence against the standard prediction

Moonlighting Game

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

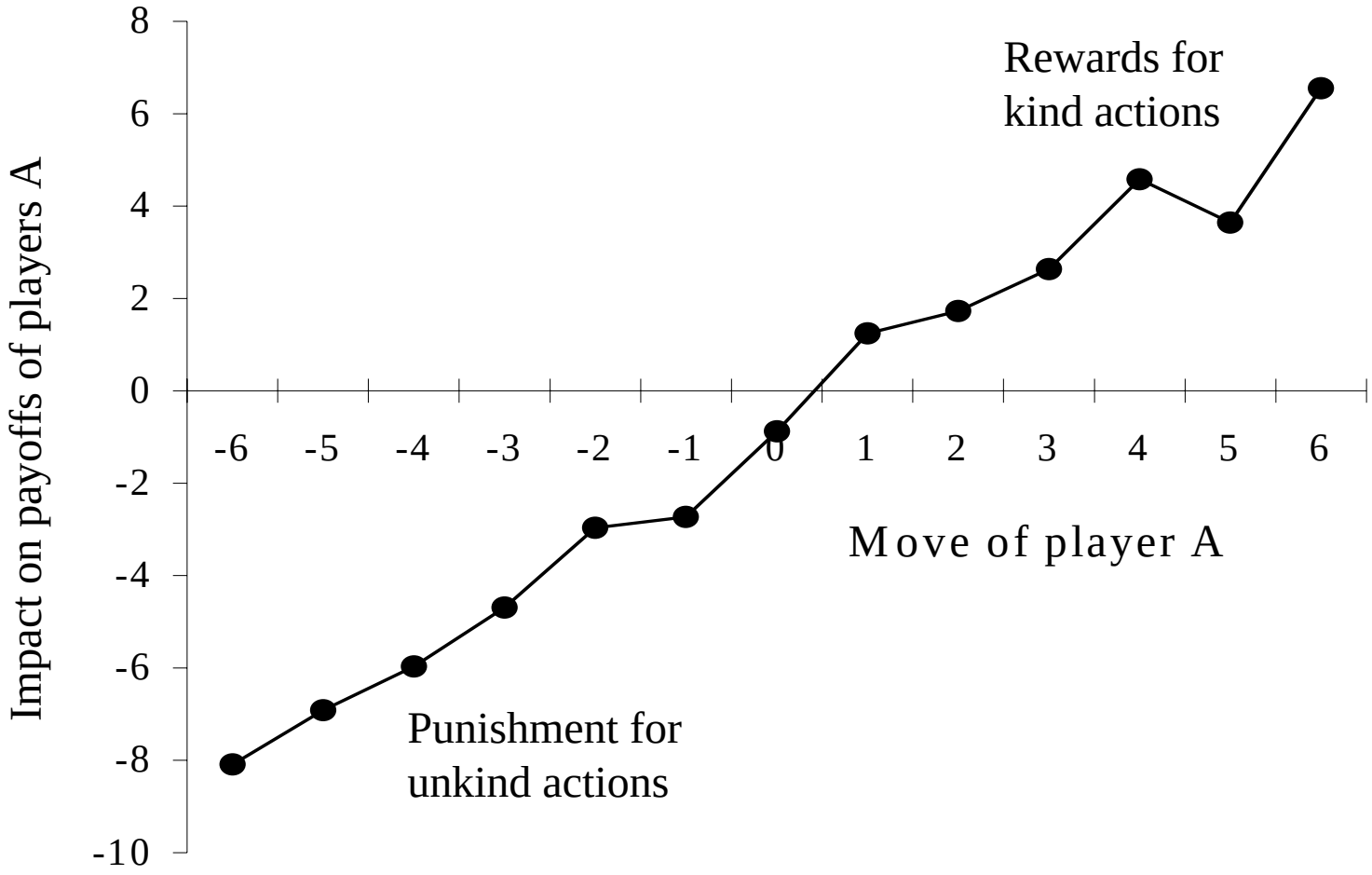
- 1. Stage:
  - Players receive an endowment of 12 points
  - Player A chooses action $a \in \{-6, -5, \ldots, 5, 6\}$
  - $a \geq 0$: A gives B a points
  - $a < 0$: A takes $|a|$ points from B
  - In case $a \geq 0$ the experimenter triples a such that B receives $3a$.
  - If $a < 0$ player A takes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

Moonlighting Game (ii)

- 2. Stage
  - B observes *a* and 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 by $3|b|$
- Standard Prediction:
  - $b = 0$ for all a, because punish/reward is costly
  - $a = -6$

Results

Falk et al. "Testing theories of Fairness, Intentions matter"



Fairness in markets? The gift exchange game

- So far fairness and reciprocity demonstrated only in bilateral or multilateral bargaining environments: Also relevant in markets?
- The impact of reciprocity on the market outcome crucially depends on whether contracts are complete or incomplete.
- **Gift-exchange game** (Fehr and Falk JPE 1999)
- Stage 1: Firms and workers enter into wage contracts, through a double auction market, with wage $\in$ [20, 120]
- There is an excess supply of workers (7:11)
- Unemployment benefit = 20

- Stage 2: Workers who agreed to a contract choose effort $\in$ [0.1, 1] , with increasing MC of effort.
- Payoffs:
  - Firms: (120 – wage)*effort
  - Workers: wage – cost of effort

e	.1	.2	.3	.4	.5	.6	.7	.8	.9	1
c(e)	0	1	2	4	6	8	10	12	15	18

- Control treatment with complete contract (effort exogenously set = .1)

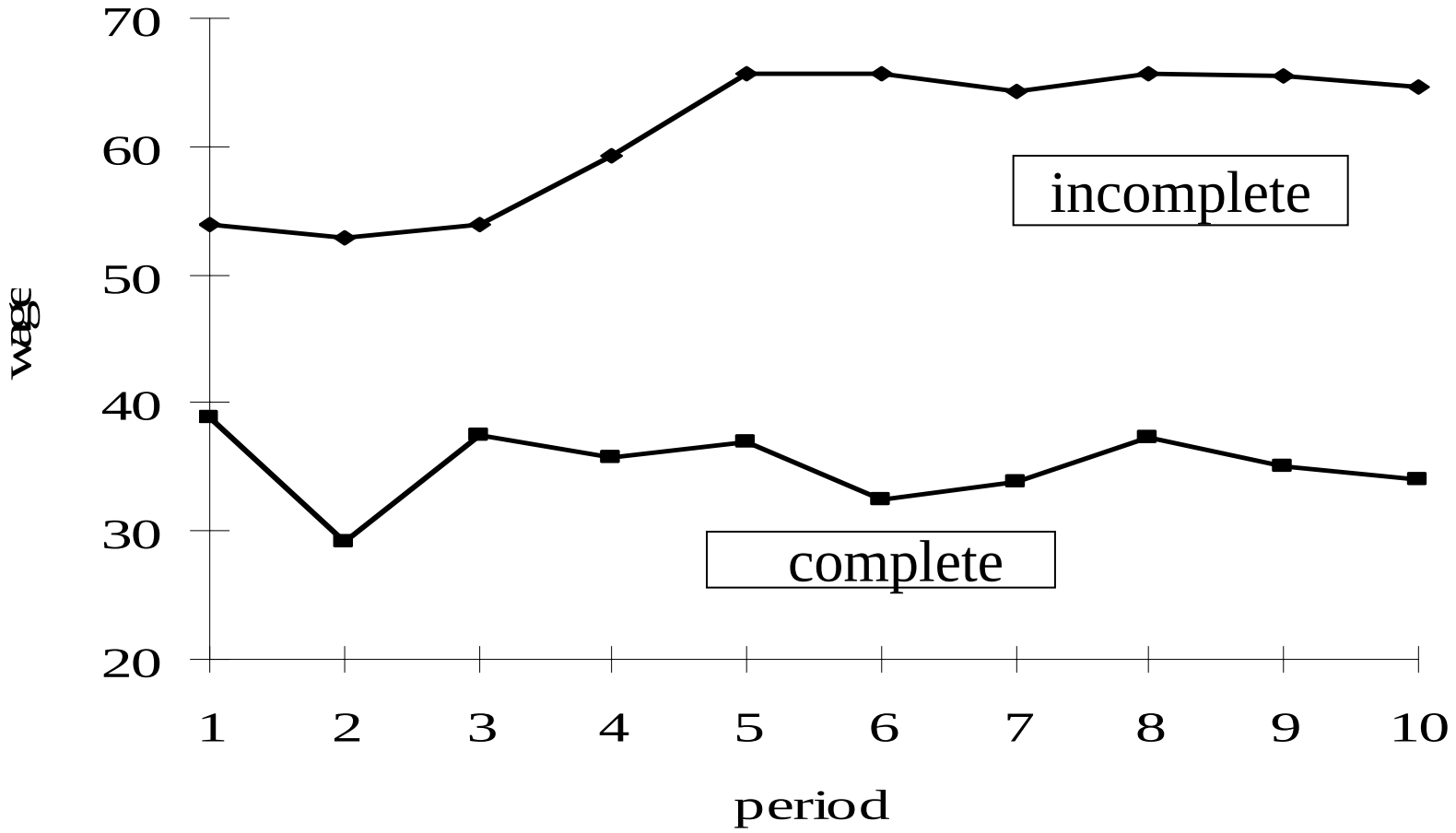
  Payoffs:

  - Firms: 120 – wage
  - Workers: wage – 20

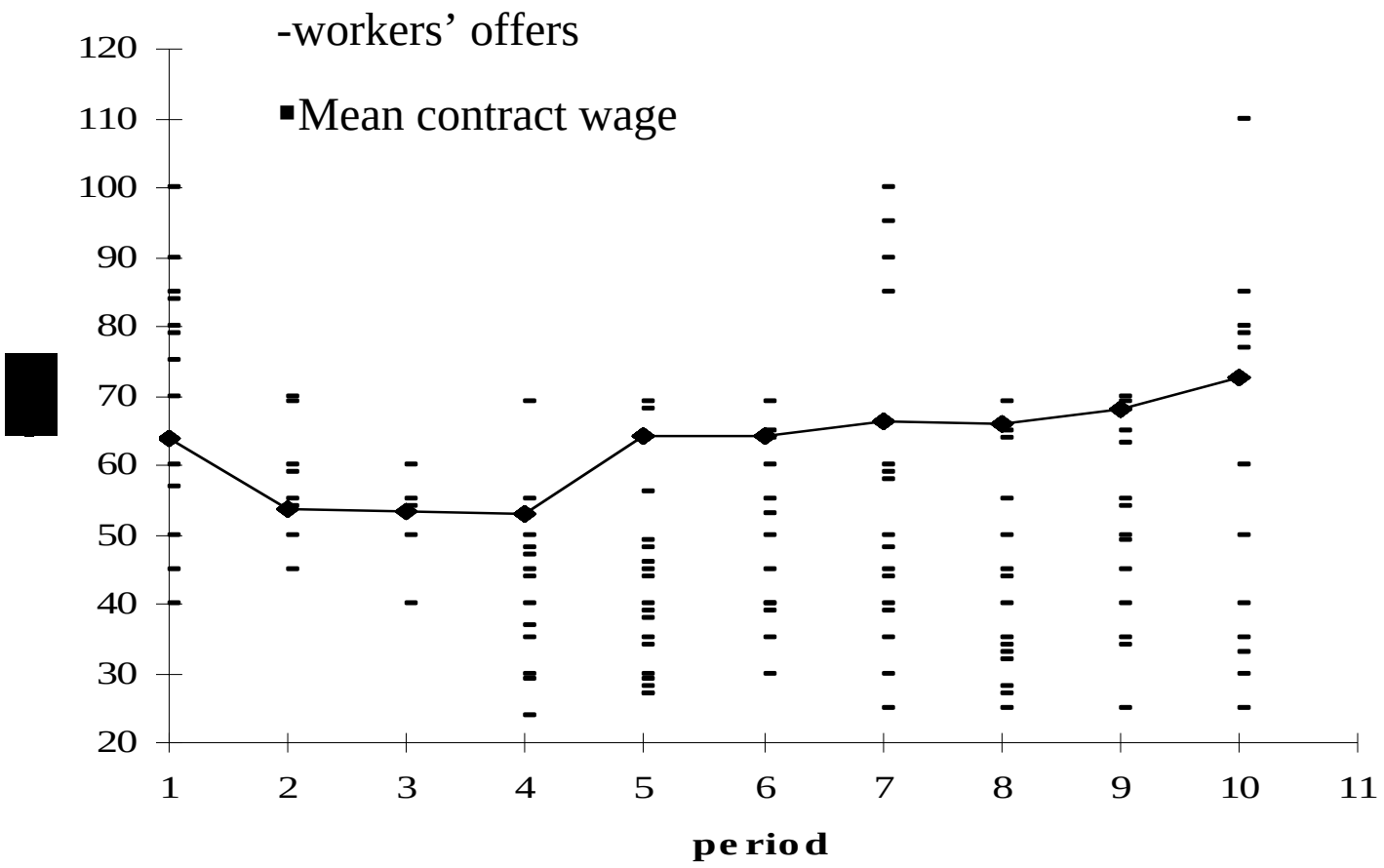
- Standard prediction:
  - wage = 20 in both treatments.
  - effort = 0.1 in the incomplete contract market

Competitive Prediction

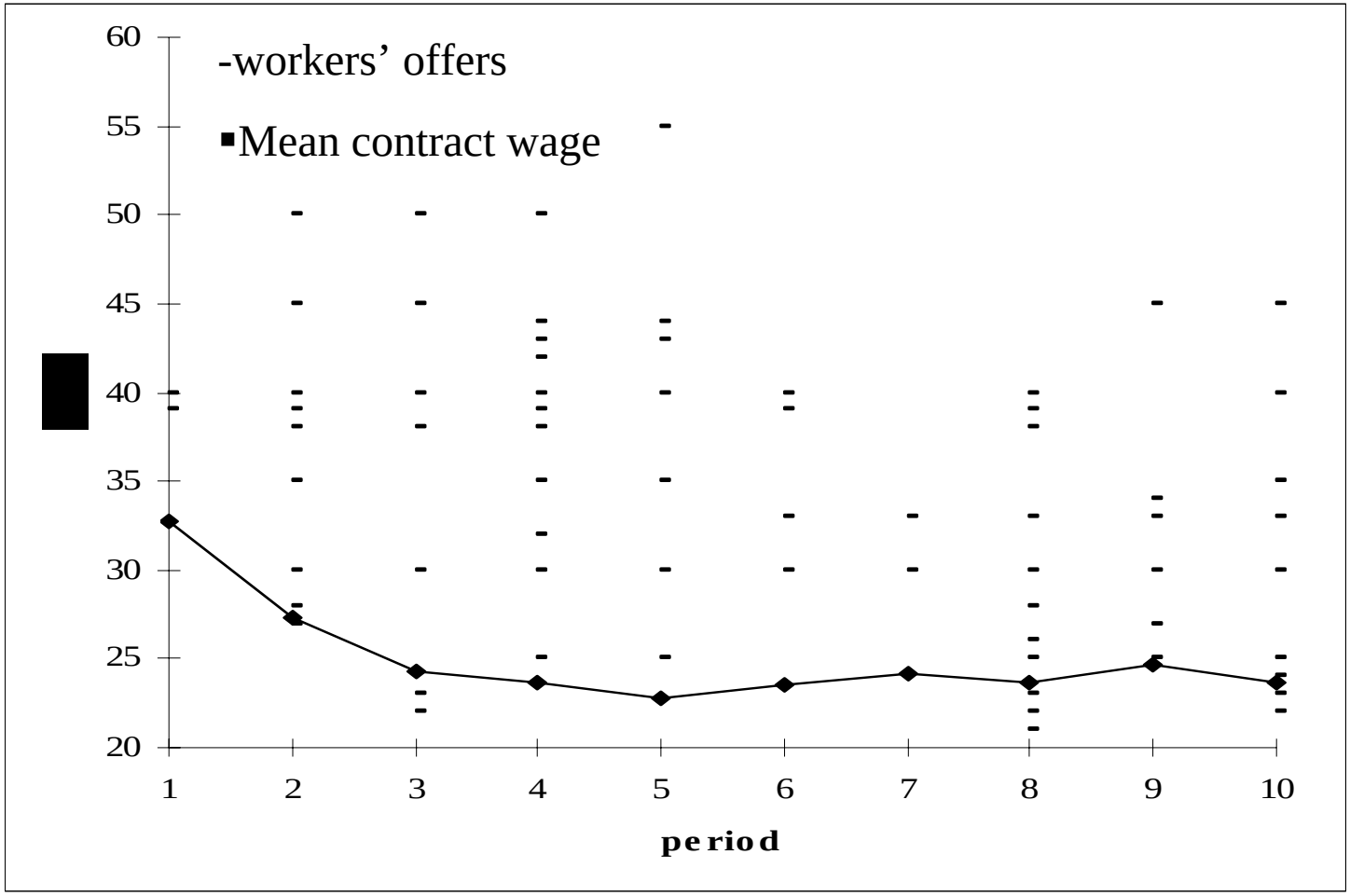
Reciprocity in Markets: Wages



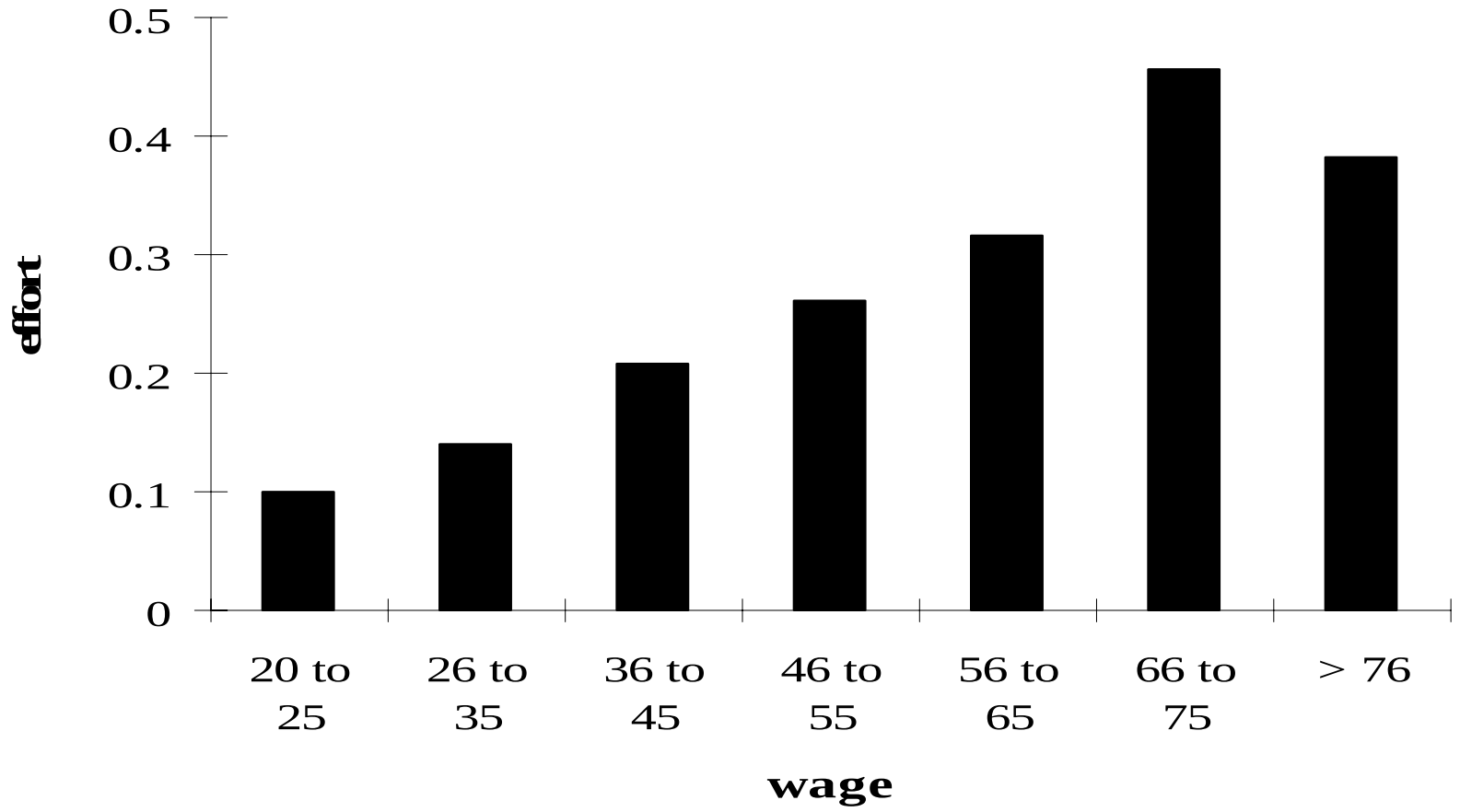
Underbidding: Incomplete Market



Underbidding: Complete Market



Reciprocity in Markets: Wage-effort Relation



Markets: Summary

- 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.

- There are many variants of this game (e.g., Fehr, Kirchsteiger and Riedl QJE 1993 (one sided auction), Gächter and Falk SJE 2002 (bilateral)
- Always the same main results

Field study

- Three treatments:
  - No gift
  - Small gift (postcard painted by children)
  - Large gift (set of four postcards painted by 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

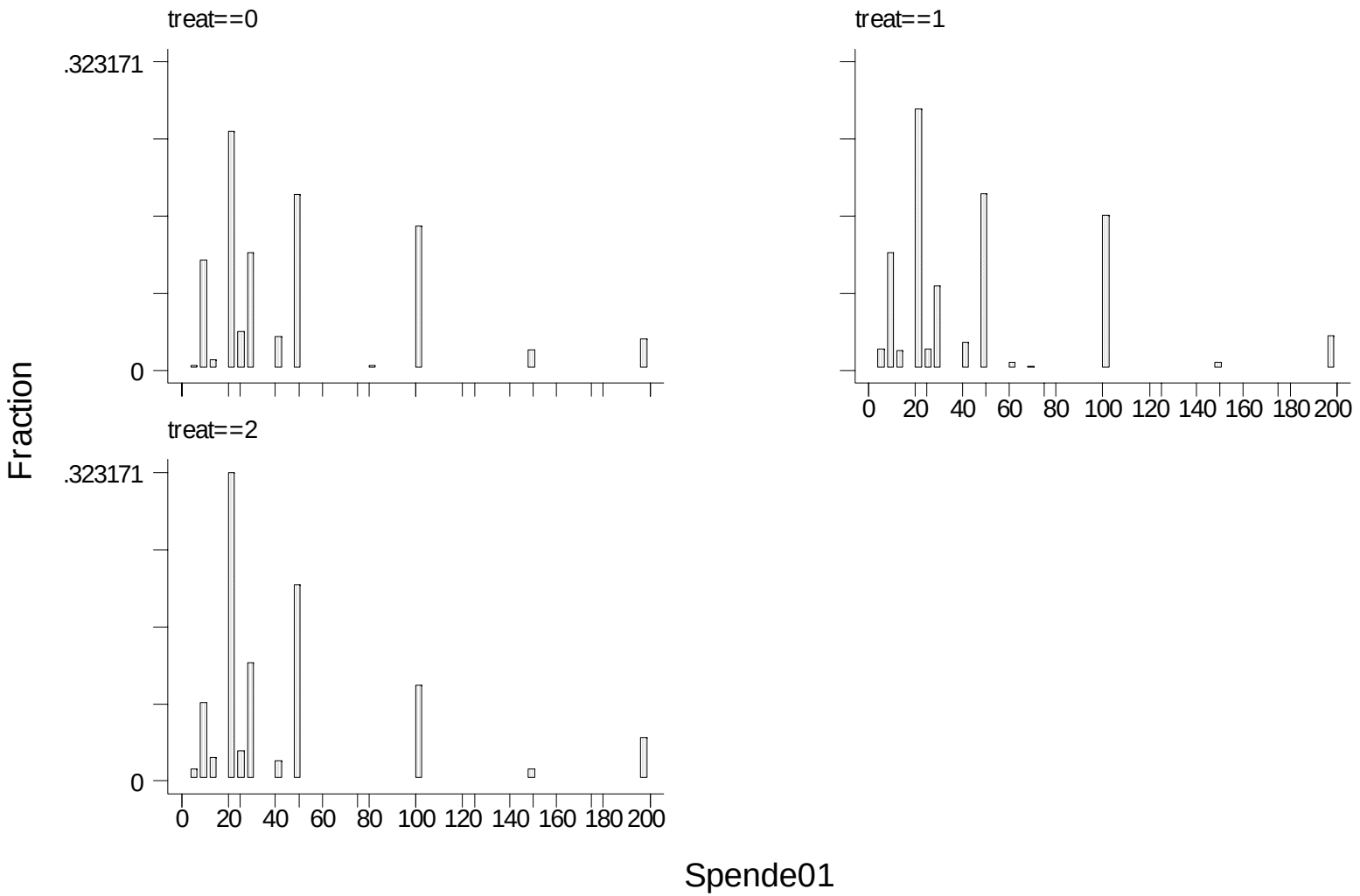
Treatment differences in the frequency of donations

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.



Histograms by treat

What Are the Puzzles ?

Game	Puzzle	No Puzzle = compatible with standard theory
Ultimatum game	Rejection of positive offers	high offers
Trust game	Payback of second movers	“Investments”
Gift exchange game	High efforts for high wages	high wages
Public goods game with punishment opportunity	Punishment	Increase of cooperation compared to treatment with no punishment opportunity
Cooperation in prisoner‘s dilemma	In one-shot game and in last period	In (infinitely) repeated games

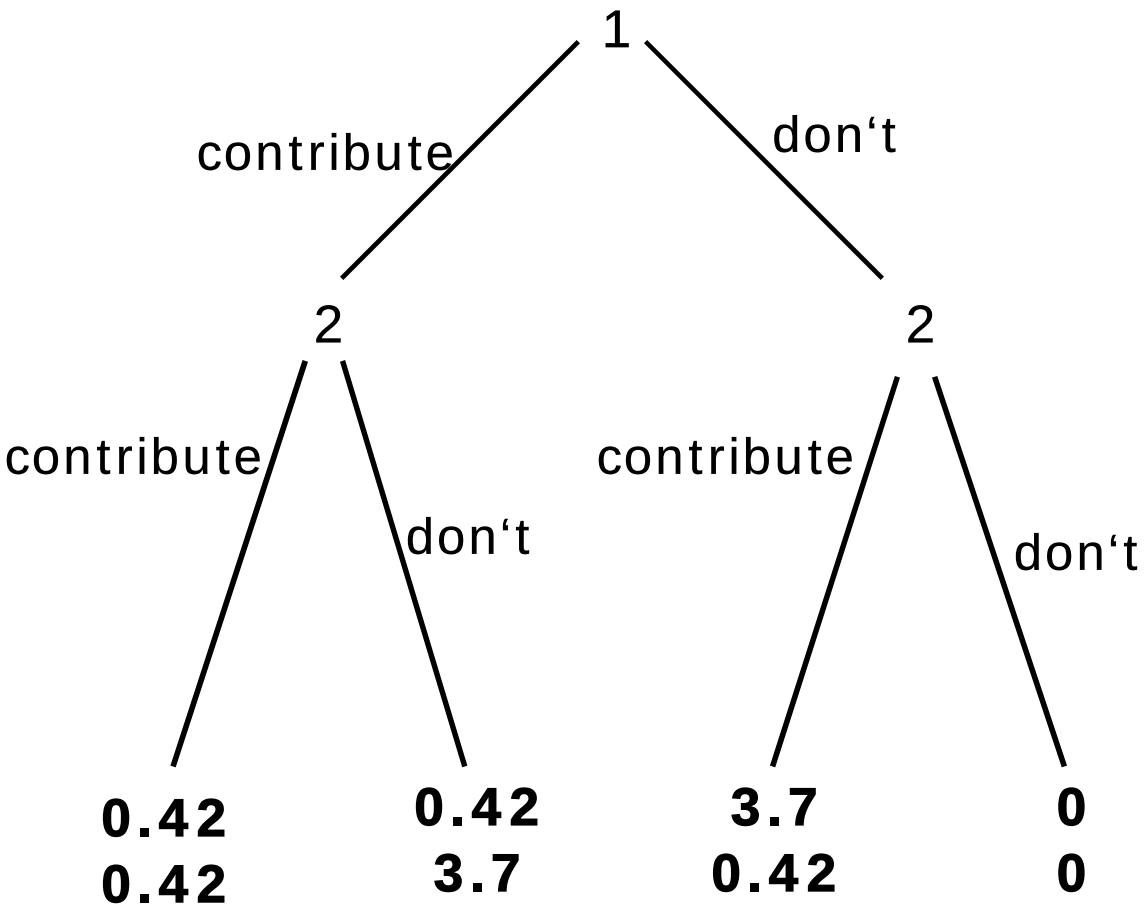
Evidence Against Fairness (?)

- In many games the experimental outcome is not in contradiction to standard theory.

  - Convergence to standard prediction in public goods games.
  - Unequal outcomes in complete, competitive markets (e.g. double auction or proposer competition).
  - Very unequal outcome in best-shot game.

Best-shot Game (Mini Version)

Harrison, Hirshleifer (1989), Prasnikar/Roth (1992)



- Players 2 accept unequal outcome of (3.7, 0.42)
- Such distributions are rarely accepted in the ultimatum game.

Understanding Fairness

- **Predictive models of fair behavior**
  - Preference based
- **How predictive models can be used**
  - They formalize intuitive ideas and make them testable.
    - Detect and distinguish between features.
  - Provide precise predictions for applications.
  - Give framework for evolutionary models.

  - Therefore, models
    - should be applicable to any game.
    - should have a constant parameter set.

The Models

- **Outcome oriented models**
    - Fehr and Schmidt (1999) (FS)
    - Bolton and Ockenfels (2000) (BO)
- **Reciprocity models** [Rabin (1993)]
    - Falk and Fischbacher (1999) (FF)
    - Dufwenberg and Kirchsteiger (1998) (DK)
- [Charness and Rabin (2002)] (CR)
- [Levine (1999)]

Outcome Oriented Models

- $U_i = U_i(\pi_i, \pi_{-i})$
  - Utility depends on own and others' payoffs

How does U_i depend on π_{-i}

- Share: $\pi_i/\Sigma\pi_j$ (BO)
- all differences: $\pi_i-\pi_j$ (FS)

The Model of Fehr and Schmidt (1999)

$$U_i = \pi_i - (\alpha_i/(n-1)) \sum \max(\pi_j - \pi_i, 0) - (\beta_i/(n-1)) \sum \max(\pi_i - \pi_j, 0)$$

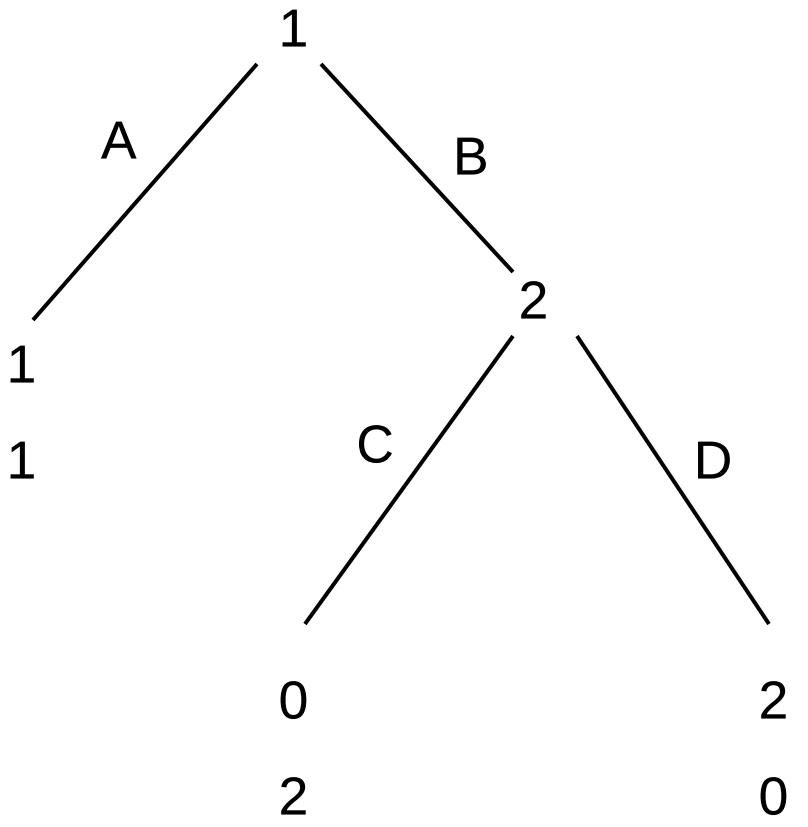
α_i = individual parameter of negative inequality aversion.

β_i = individual parameter of positive inequality aversion.

For selfish subjects, both parameters are equal to 0.

- Assumptions: $\alpha_i \geq \beta_i \geq 0$, $\beta_i < 1$
  - Negative inequality aversion is more important than positive.
  - Nobody destroys money to reduce positive inequality.

A Game Tree and Psychological Game Theory



Is player 1‘s move B kind or unkind?

- Kindness depends on what player 1 expects player 2 will choose.
- (second order) beliefs are important

- Concept of Psychological Game Theory: Geanakoplos, Pearce, Stacchetti (1989)
- Example by Dufwenberg and Kirchsteiger

Reciprocity Models (Rabin,DK,FF)

- The structure of reciprocity models

$$U_i = \pi_i + \rho_i \sum \text{kindness}_{\ j \text{ to } i} * \pi_j$$

- What determines kindness
  - Payoff that player i is supposed to receive compared to reference payoff
    - Absolute reference (FF)
    - Relative reference (DK)
  - Intentions are incorporated into the theory by considering the alternatives.

  - Levine model (and CR [extended version]): kindness depend on the type of the other player.

Kindness in the FF-model

- Kindness term =
    - Outcome term weighted with intention factor
- Outcome term = $\pi_i - \pi_j$
- Intention factor depends on alternatives
    - An action is **intentionally kind** if the other player had any alternative to be less generous (give me less).
    - An action is **intentionally unkind** if the other player had a *reasonable* alternative to be more generous.
- The model combines the equity standard of the outcome oriented models with an intention concept similar to the other reciprocity models.

A Questionnaire on kindness (Falk and Fischbacher 2001)

- Subjects get list of possible offers.
- Have to evaluate the kindness of these offers.
- Between –100 and +100.

(π_j, π_i)	i	ii	iii	iv	v	vi	vii	viii	ix
(0, 10)	72.3					79.9	73.4		80.3
(1, 9)	68					73.3	62		72.5
(2, 8)	62	75.3		41.1	61.2	61.9	40.8		62.2
(3, 7)	51.4								
(4, 6)	40								
(5, 5)	29.4	33.4							27.9
(6, 4)	-23.2								
(7, 3)	-52.9								
(8, 2)	-71.9	-70.6	-31.5		-47.7	-50.5		-9.1	-60.9
(9, 1)	-84.5					-80.3		-56.4	-82.6
(10, 0)	-95.4					-97.3		-88.8	-97.3

Questionnaire on Kindness

Summary of four main conclusions:

1. If *j* has only one option, kindness and unkindness are much weaker.
2. Even if *j* has only one option, kindness and unkindness not zero. Outcomes matter.
3. A kind offer is viewed as similarly kind across games, as long as *j* could have made less kind offers. Vice versa for unkind.
4. Action is not so unkind if *j* had no “reasonable” alternative, I.e., alternative that is not too harmful to *j*.

Distinctive Feature of the Models

- Differences in preference for distributions.
    - Relative importance of inequity aversion, efficiency and maximin preferences.
- Differences in reciprocal behavior
    - Who is the relevant reference agent?
        - ➢individual or group
    - Is all punishment driven by inequity aversion?
        - ➢difference reduction or retaliation
    - What is the role of intentions?
        - ➢outcomes or intentions

Differences in Reciprocal Behavior Falk/Fehr/Fischbacher (2000)

- Q1: Who is the relevant reference agent?
  - ➢ individual or group (FS or BO?)
- Q2: Is all punishment driven by inequity aversion?
  - ➢ difference reduction (inequity aversion) or retaliation (reciprocity)
- Q3: What is the role of intentions?
  - ➢ outcomes or intentions

Q1: Who Is the Relevant Reference Agent?

- Three person one-shot public goods game with punishment opportunity:
- 1st Stage: public goods game
  - Contribute 20 points (cooperate) or nothing (defect)
  - Payoff

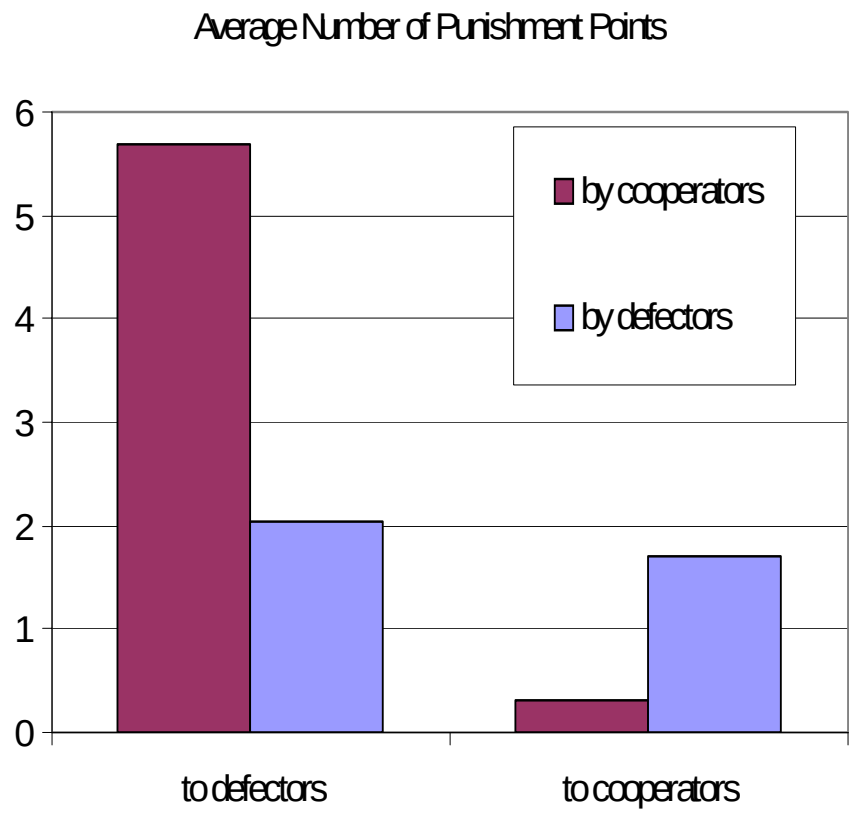
    20 - own contribution +

    0.6 * sum of all contributions
- 2nd stage: Reduce the other player's payoff at a cost
  - Punishing cooperators: 1 point reduction costs .3 points.
  - Punishing defectors: 1 point reduction costs .4 points.
  - It is cheaper to punish cooperators.

Q1 : Prediction

- BO predict that if cooperators punish, they punish other cooperators.
  - It is the cheapest way to reduce inequity because it reduces the average payoff of the other players most (inequity measured towards the whole group and not individually)
- The other theories predict that if cooperators punish, they punish defectors.
  - Because fairness is evaluated for each other player separately, those are punished who “deserve” punishment. Either because they have a higher payoff (FS) or because they are unkind (DK and FF).

Q1 : Experimental Result (N=120)



61 percent of the subjects cooperate.

From the cooperators:

- 69 percent punish
- 67 percent punish defectors only
- 7 percent punish cooperators only

From the defectors (39 percent)

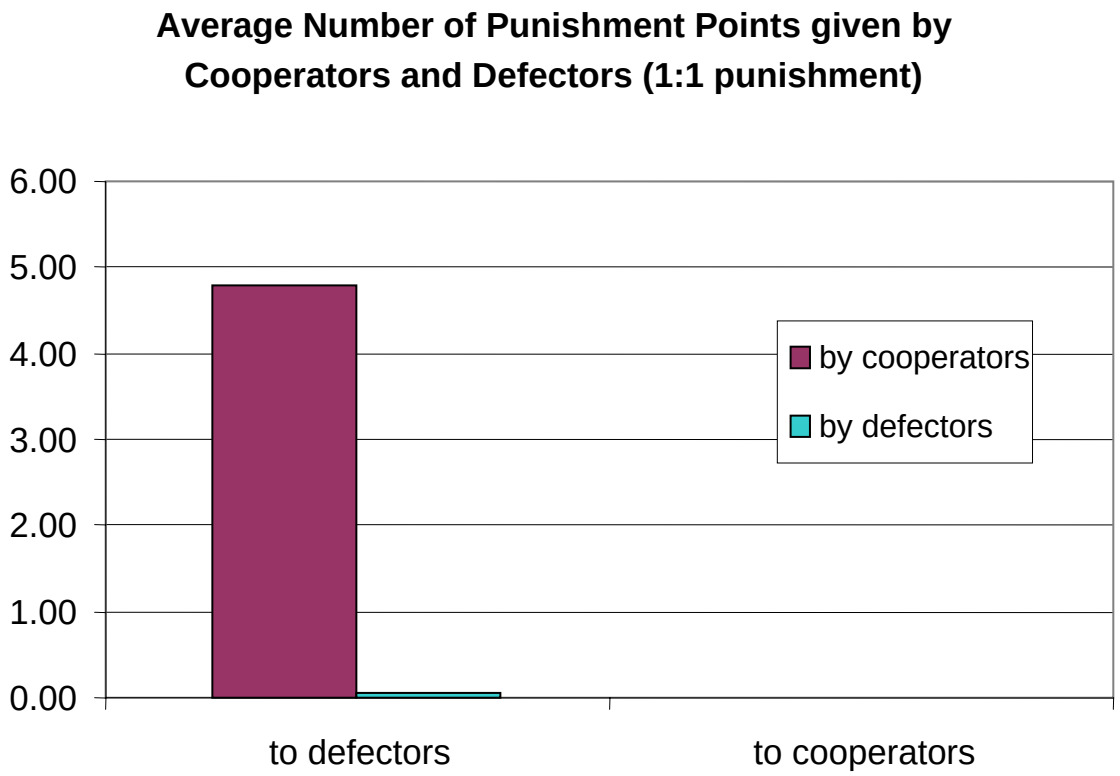
- 49 percent punish, cooperators and defectors

Q2: Is All Punishment Driven by Inequity Aversion?

- One-to-one punishment
  - Same three person one-shot public goods game with punishment opportunity, but
  - 1 point reduction costs 1 points, i.e., there are *higher* costs of punishment

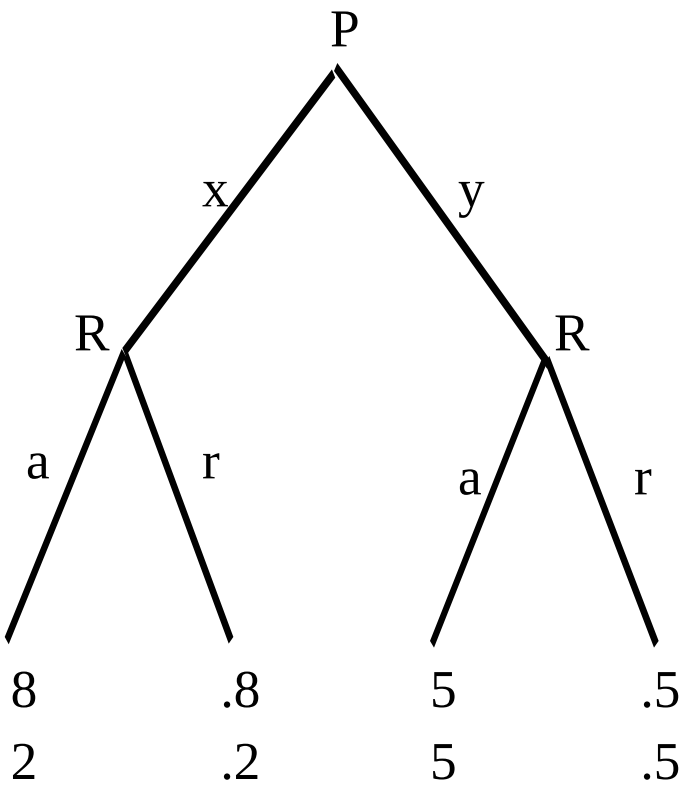
- Inequity aversion models predict no punishment because inequity cannot be reduced.

Q2: Is All Punishment Driven by Inequity Aversion?



- 51 percent cooperate
- of these cooperators 47 percent punish defectors
- punishment behavior is incompatible with any equity model
- Defectors do not punish.

Q2: Is All Punishment Driven by Inequity Aversion?

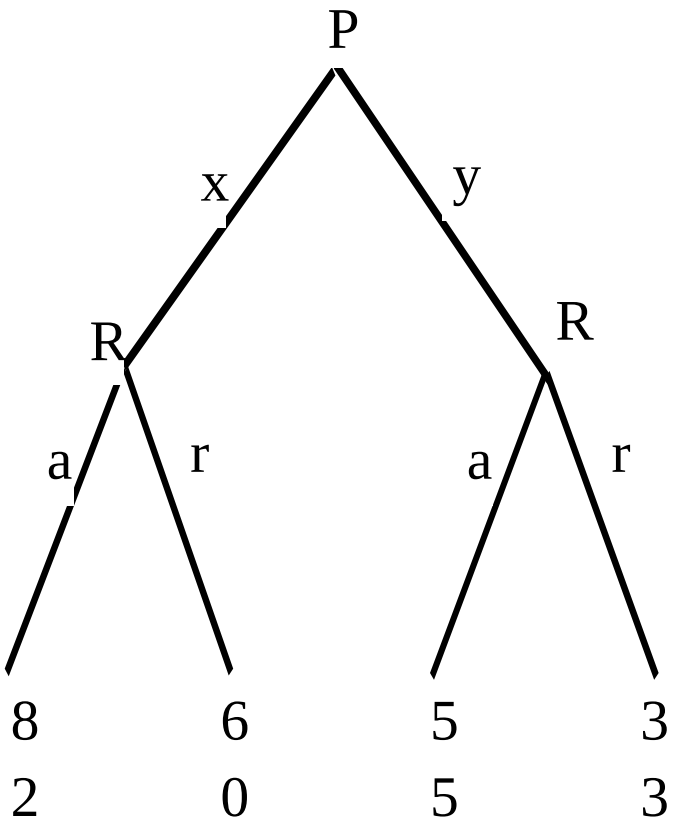


UG with constant relative share

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

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

Exp 2: Is All 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.

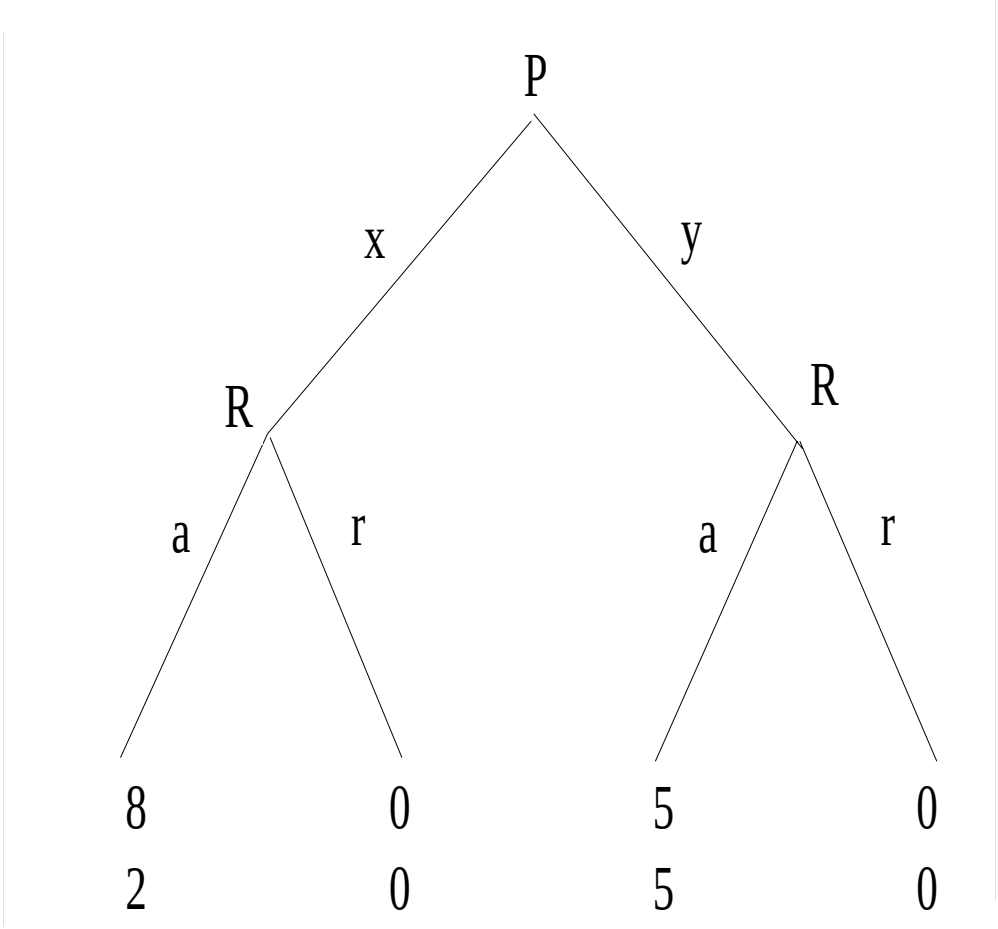
Q2: Experimental Results (N=48)

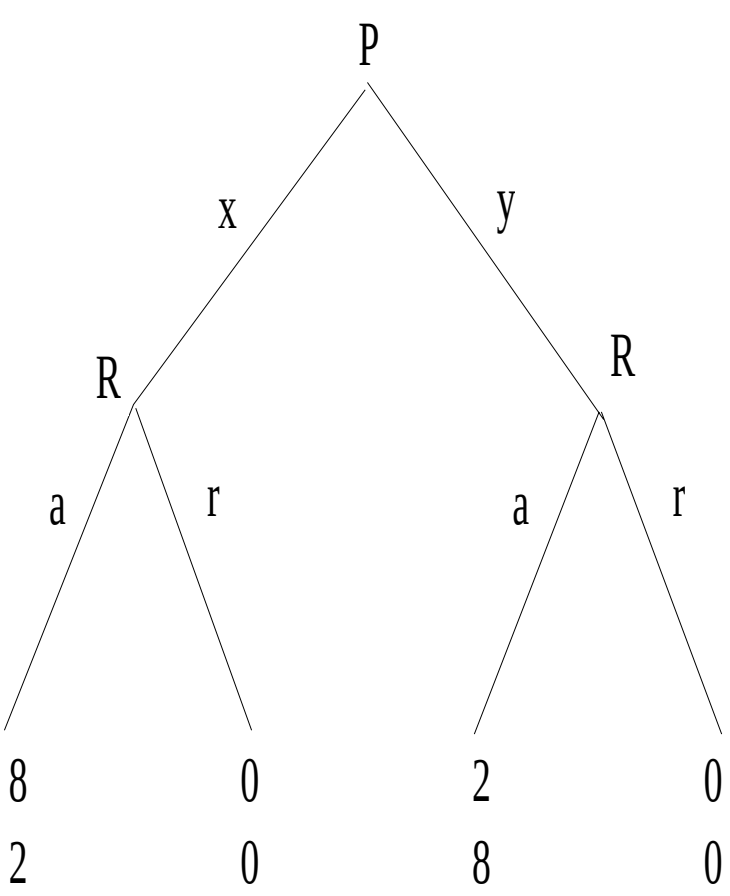
	Predict rejections of 8:2 offer				Result Rejection rate
	BO	FS	DK	FF	
UG with constant relative share	no	yes	yes	yes	38%
UG with constant difference	no	no	yes	yes	19%

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

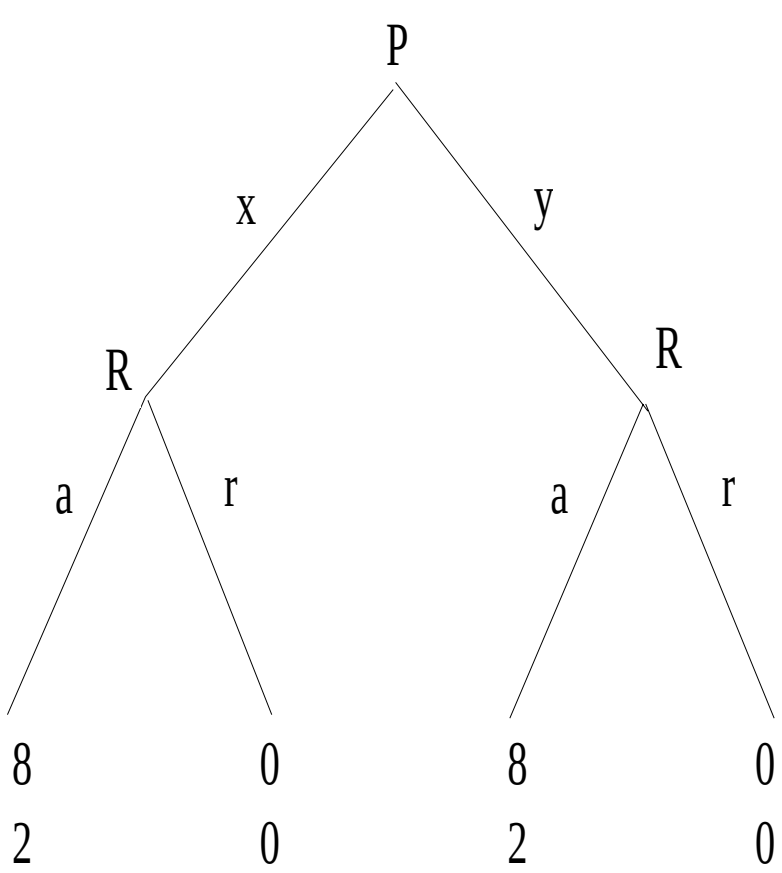
Q3: Are Intentions Important? Four Mini Ultimatum Games

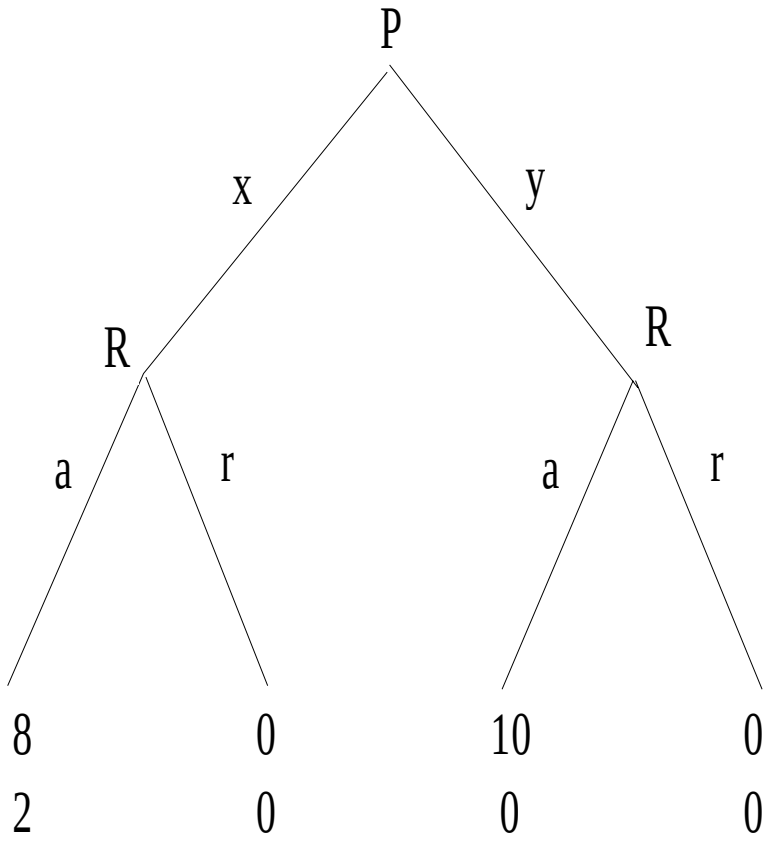
This is like the best shot game





Intentions (ii)



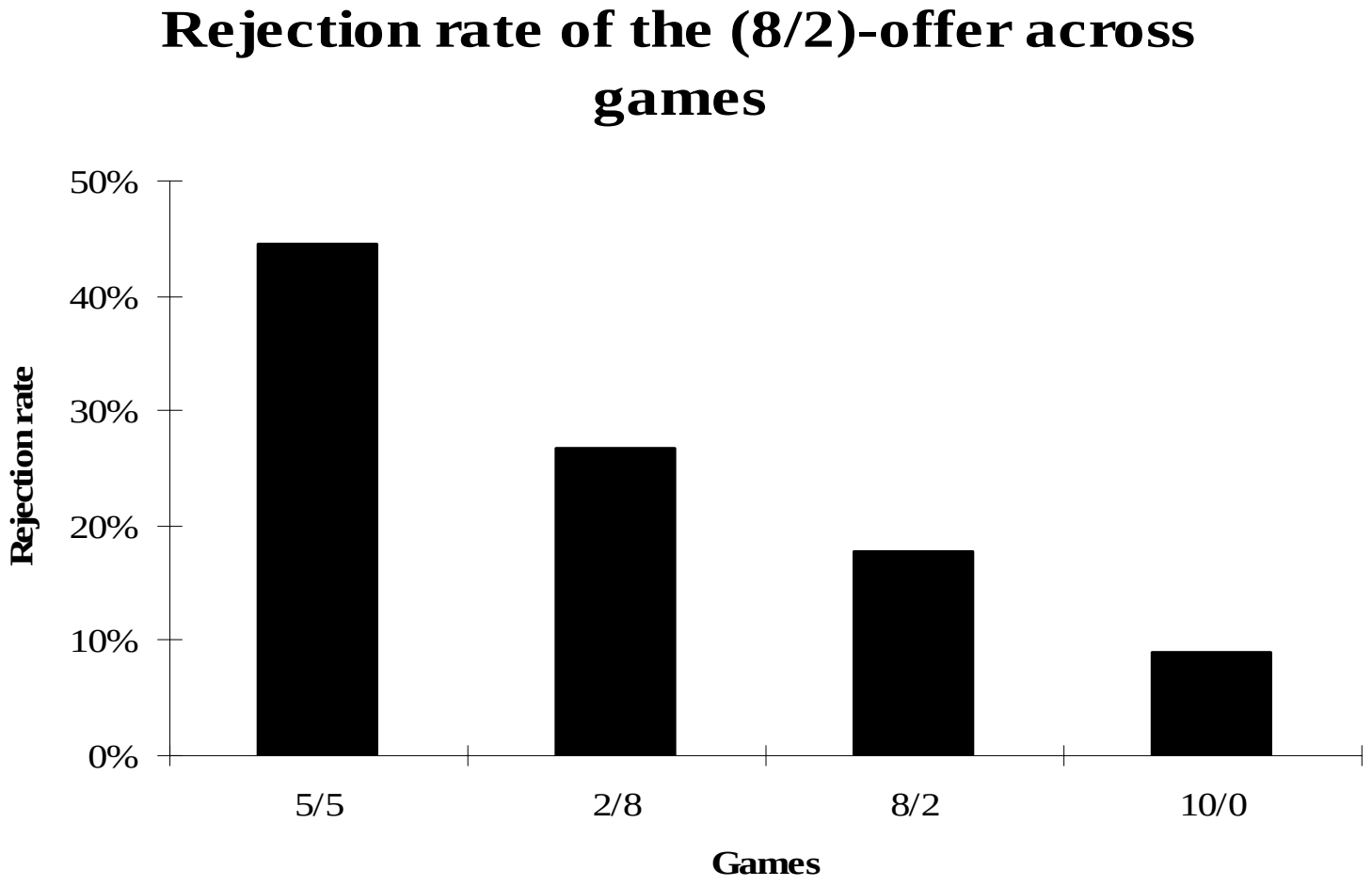


Predictions of the rejection rates of the 8:2 offer in left-hand subgame

- BO and FS predict the same rejection rate regardless of the alternatives.
  - These theories model fairness in a consequentialist way and the consequence of the 8:2 offer is always the same.
- DK predict zero rejection rate for the alternative 10:0.
  - Subjects do not consider the 8:2 offer as unkind because 10:0 is even more unkind.
- FF predict positive rejection rates in all cases. The rejection rate is highest in the 5:5 case.
  - Fairness is determined by the outcome *and* the intention of the other subject, so 8:2 can be rejected even when alternative was worse, due to outcome concerns.

Experimental results (N=45)

Falk, Fehr and Fischbacher, Economic Inquiry forthcoming)



Q3: Proposer Behavior

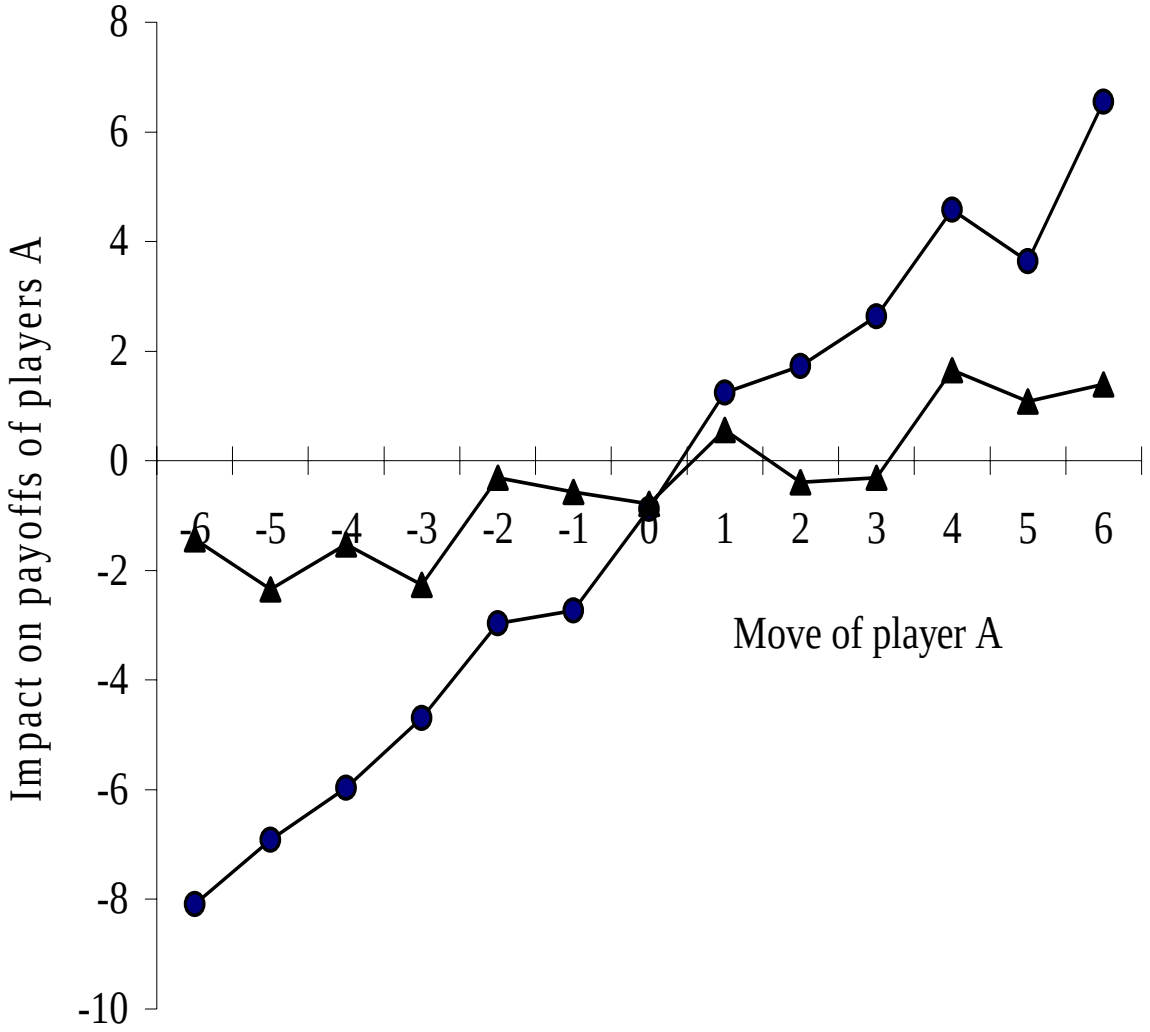
Alternative	Rejection rate of 8/2	Rejection rate of alternative	Choice of 8/2
5/5	44%	0%	31%
2/8	27%	2%	73%
8/2	16%	20%	-
10/0	9%	89%	100%

- Proposer behavior is compatible with selfishness, but also with preferences for fairness.

Intentions, once more

- Moonlighting game as before (see above).
- 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.

Intentions and Random Move Games

- If the move of the first player in the ultimatum game is made by a random device, then
  - An unfair outcome is not intended by player 1.
  - Therefore, unfair offers are less likely to be rejected.

  (Blount 1995)
- Same idea in gift exchange game (Charness)
  - Here, high wages are rewarded with similar effort in the treatment in which a person chooses the wage compared to the treatment in which the wage is randomly drawn. (though steeper slope)
- In both experiments reward and punishment also occur in the random move treatments.

Conclusion

- Fairness can be captured by incorporating preferences for fairness into the utility function.
- Models reconcile results with equal outcomes as well as with unequal outcomes (e.g. UG vs. competitive markets).

- Fairness is evaluated individually.
- Inequity reduction is not the only reason for punishment.
- Intentions and outcome matter.

- Reciprocity models give a better description of human behavior – but at a cost in tractability.