AREC 815: Experimental and Behavioral Economics

Reciprocity and Conditional Cooperation

Professor: Pamela Jakiela

Department of Agricultural and Resource Economics
University of Maryland, College Park

Conditional Cooperation in Public Goods Games

Defining Reciprocity

An agent has reciprocal preferences if the utility weight on others depends on whether they are kind/mean, cooperative/uncooperative, etc.

- **Positive reciprocity:** the willingness to reduce one's own payoff to reward those who have been kind, helpful, cooperative, etc.
- **Negative reciprocity:** the willingness to reduce one's own payoff to punish those who have been unkind, unhelpful, uncooperative, etc.

When might reciprocity be economically important?

- Voluntary public goods provision
- The labor market (shirking, bonuses)
- Decentralized markets involving trust
- Other settings?

Public Goods Games

Each subject i divides her endowment, e_i, between a private account and a shared account that is evenly divided among all group members

- Contributions to shared account multiplied by factor $m > 1$
- Individual payoff: $\pi_i = e_i - g_i + m \left(\sum_n g_n \right) / N$
- Optimal strategy: **free-riding**

Conditional Cooperation in Public Goods Games

Fact #1: both conditional cooperators and free riders are common

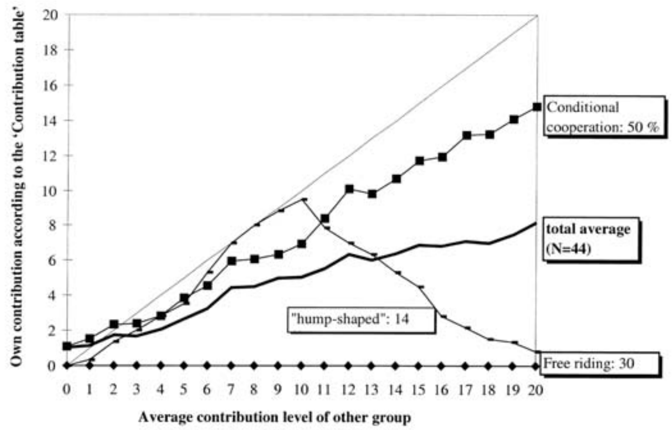
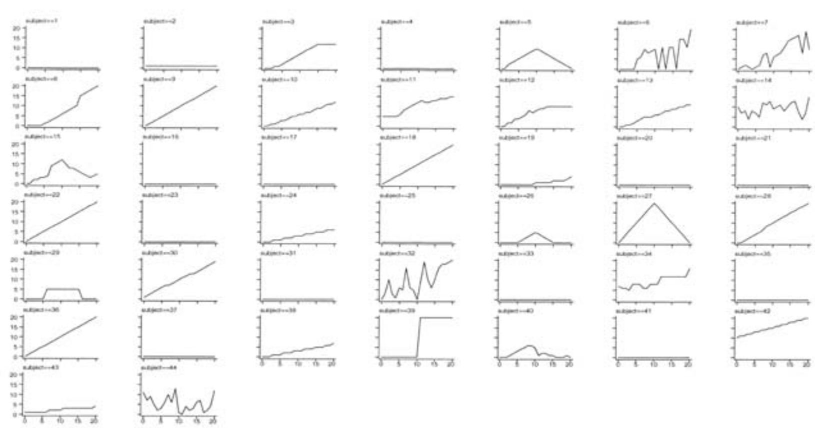


Fig. 1. Average own contribution level for each average contribution level of other members (diagonal = perfect conditional).

Source: Fischbacher *et al* (2001)

Conditional Cooperation in Public Goods Games

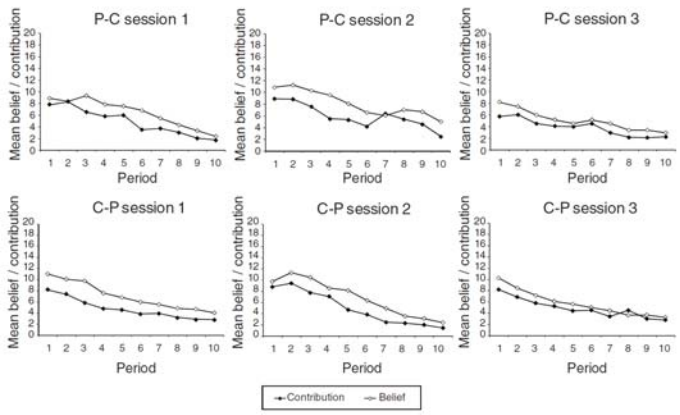
Individual contribution schedules: $\tilde{g}_i\left(\sum g_{-i}/(N-1)\right)$



Source: Fischbacher *et al* (2001)

Conditional Cooperation in Public Goods Games

Fact #2: contributions decline over time; beliefs do, too



Source: Fischbacher and Gächter (2010)

What Explains Beliefs?

How do Fischbacher and Gächter "model" beliefs?

What Explains Beliefs?

	Dependent variable: Belief about other group members' contribution		
Model	(1)	(2)	(3)
Period	−0.761***	−0.079	
	(0.090)	(0.042)	
Others' contributions ($t-1$)		0.394***	0.415***
		(0.023)	(0.020)
Belief ($t-1$)		0.549***	0.569***
		(0.037)	(0.036)
Constant	10.711***	0.835*	0.118
	(0.864)	(0.398)	(0.148)
Observations	1,260	1,260	1,260
R^2	0.26	0.64	0.64

Notes: OLS regressions with data from period 2 to 10. Robust standard errors (clustered on sessions) in parentheses.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Source: Fischbacher and Gächter (2010)

Predicting Contributions

Data from the P-experiment can be used to predict contributions:

- $\tilde{g}_i\left(\sum g_{-i}/(N-1)\right)$ is i's **conditional contribution**
- Given belief $\tilde{b}_{i,t}$, $\tilde{g}_i\left(\tilde{b}_{i,t}\right)$ is i's **predicted contribution**

Estimate OLS regression specification:

$$Contribution_{i,t} = \alpha + \gamma Period_t + \lambda PredictedContribution_{i,t} + \delta Belief_{i,t} + \varepsilon_{i,t}$$

Explaining the Decline in Cooperation

Model	Dependent variable: Contribution (1)	(2)	(3)	(4a)	(4b)	(4c)
Periods used	1–10	1–10	1–10	1–10	1–5	6–10
Subjects excluded[a]	No	No	No	Yes	Yes	Yes
Period	−0.639 (0.071)***	−0.060 (0.056)				
Predicted contribution		0.242 (0.069)**	0.242 (0.069)**	0.443 (0.073)***	0.385 (0.074)***	0.614 (0.082)***
Belief		0.644 (0.071)***	0.666 (0.059)***	0.545 (0.065)***	0.582 (0.065)***	0.376 (0.116)**
Constant	8.343 (0.545)***	0.005 (0.569)	−0.473 (0.244)	−0.318 (0.312)	−0.204 (0.541)	−0.116 (0.378)
Observations	1,400	1,400	1,400	1,260	630	630
R^2	0.10	0.34	0.34	0.38	0.33	0.33

Note: Robust standard errors in parentheses.
[a] Models 4a to 4c exclude (confused) subjects who, on the basis of the P-experiment, could not be classified according to the FGF classification as either a "free rider," "conditional cooperator," or a "triangle contributor."
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Source: Fischbacher and Gächter (2010)

Discussion

What do we conclude from this analysis?

- Beliefs matter, over and above predicted contributions
  - Are subjects investing in goodwill early on?
- Conditional cooperation is less than one-for-one
  - Necessarily declines over time?

What is missing from this analysis?

Reciprocity in the Labor Market

Reciprocity in Experimental Labor Markets

Fehr et al (1993) propose simple labor market experiment

- Subjects randomly chosen to be "firms" offer those chosen to be"workers" a wage, w, and stipulate a desired level of effort
- Firms cannot punish employee deviations from agreed effort level

Fair wage-effort hypothesis: labor relations as a gift exchange game, firms may offer high wages to extract greater effort from workers

- Positive reciprocity $\Rightarrow$ workers reward above market-clearing wages
- Negative reciprocity $\Rightarrow$ workers punish "unfairly" low wages
- Employers may reward, punish deviations from expected worker effort

Reciprocity in Experimental Labor Markets

Worker's payoff, conditional on wage p_j, effort level e_j:

$$u_j = p_j - c - m(e_j)$$

$m(e)$-SCHEDULE

e	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$m(e)$	0	1	2	4	6	8	10	12	15	18

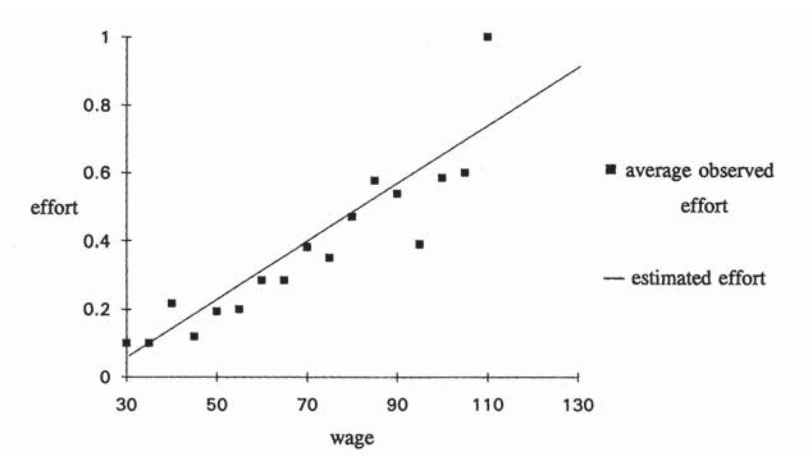
$\Rightarrow$ Workers choose minimum effort level, e_{min}

Employer's payoff, conditional on worker wage and effort:

$$\pi_i = (v - p_j)\, e_j$$

Reciprocity in Experimental Labor Markets

Result: wages, effort levels exceed equilibrium predictions



Reciprocity in Experimental Labor Markets

Fehr et al (1997) take labor market experiments further, proposing a two-stage market design that allows employers to detect "shirking"

SEQUENCE OF EVENTS

• Firms simultaneously post employment contracts $(w, \tilde{e}, f)$. • Workers observe all contracts and choose among the available offers in a randomly determined order.	Stage 1
• Workers who accepted an offer choose $e \gtreqless \tilde{e}$. • Random device determines whether shirking $(e < \tilde{e})$ is verifiable. • Firms are informed about the effort choice of their worker.	Stage 2

Let f_0 denote maximum feasible level of the penalty for shirking

⇒ Maximum enforceable effort level: $e_0 = c^{-1}(sf_0)$

⇒ Optimal contract: $(c(e_0), e_0, f_0)$

Reciprocity in Experimental Labor Markets

No reciprocity treatment (NRT):

- Workers obliged to provide effort level in contract

Weak Reciprocity Treatment (WRT):

- f_0 is low: maximum enforceable effort level is 0.1
- Firms cannot enforce high effort
- High wages may extract high effort from reciprocal workers

Strong reciprocity treatment (SRT):

- Firms choose $p \in [0, 2]$, where p multiplies worker payout
- Cost of p convex in deviations from 1

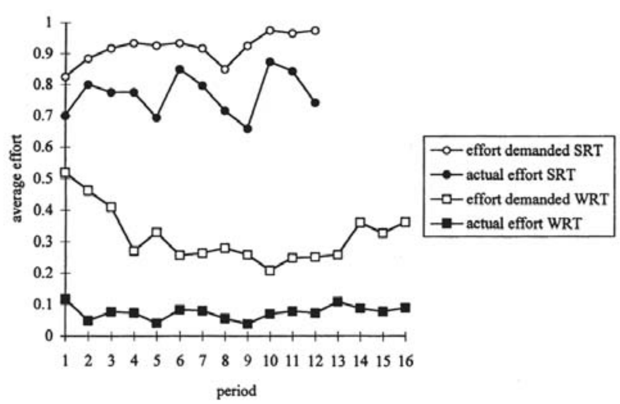
Reciprocity in Experimental Labor Markets

Testable predictions:

- Firms may offer higher rents in the WRT than in the NRT
  - Firms may hope that workers are motivated by positive reciprocity
- In the SRT, firms motivated by positive or negative reciprocity
  - Anticipating this, workers should exert greater effort

Reciprocity in Experimental Labor Markets

WRT results contrast with those of Fehr et al (1993)



Reciprocity in Experimental Labor Markets

TABLE VII

EFFORT BEHAVIOR IN THE WRT AND THE SRT

		Shirking $e < \tilde{e}$		No Shirking $e = \tilde{e}$	Excess Effort $e > \tilde{e}$	
Treatment	No. Trades	% of Trades with $e < \tilde{e}$	Average Amount of $(\tilde{e} - e)/\tilde{e}$	% of Trades with $e = \tilde{e}$	% of Trades with $e > \tilde{e}$	Average Amount of $(e - \tilde{e})/(1 - \tilde{e})$
WRT	509	65.42	0.97	33.01	1.57	0.20
SRT	144	20.83	0.82	72.22	6.94	0.83

Reciprocity in Experimental Labor Markets

TABLE VIa

FIRMS' PUNISHMENT/REWARD DECISION AT STAGE THREE, GIVEN WORKERS' EFFORT DECISION

Actual Punishment/Reward:	Shirking $e < \tilde{e}$ 30 trades	No Shirking $e = \tilde{e}$ 104 trades	Excess Effort $e > \tilde{e}$ 10 trades
$p < 1$	18 (0.19)	not possible	not possible
$p = 1$	12	52	6
$p > 1$	not possible	52 (1.62)	4 (1.53)

Note: The number in parentheses shows the average level of p.

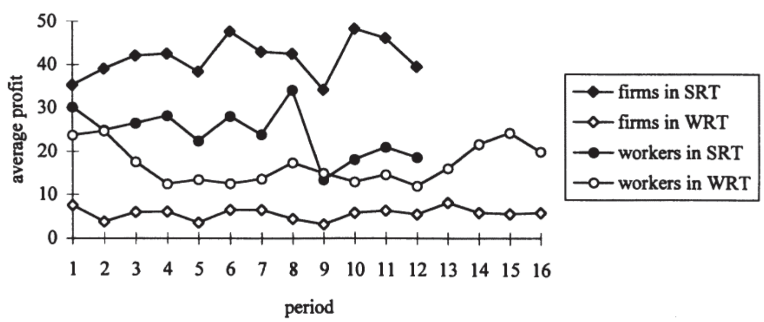
TABLE VIb

WORKERS' EXPECTATION FORMATION: DO THEY ANTICIPATE FIRMS' RECIPROCITY?

Expected Punishment/Reward:	Shirking $e < \tilde{e}$ 30 trades	No Shirking $e = \tilde{e}$ 104 trades	Excess Effort $e > \tilde{e}$ 10 trades
$p^e < 1$	18 (0.59)	not possible	not possible
$p^e = 1$	12	29	0
$p^e > 1$	not possible	75 (1.51)	10 (1.61)

Note: The number in parentheses shows the average level of p^e.

Reciprocity in Experimental Labor Markets



Reciprocity in a Field-Experimental Labor Market

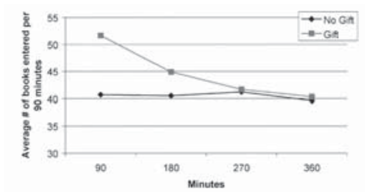
Gneezy & List (2006): field experiment with short-term employees completing (non-experimental) effort tasks for hourly wages

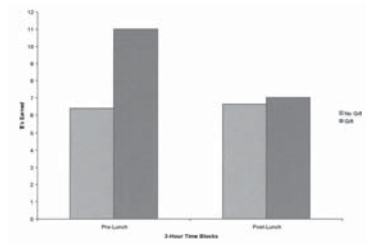
- Effort tasks: library data entry, charity solicitations

Recruitment materials employees quoted a specific hourly wage rate

Gift treatment: higher wage announced once workers arrive at job

Reciprocity in a Field-Experimental Labor Market





AVERAGE EARNINGS WITHIN TREATMENT BY 3-HOUR BLOCK*

	Prelunch	Postlunch	Difference
Gift exchange	11.00	7.026	3.974**
	(1.443)	(0.786)	(1.66)
Nongift exchange	6.40	6.633	−0.233
	(1.803)	(1.389)	(1.29)

Negative Reciprocity: Labor Disputes

FIGURE I
Month-by-Month Comparison of Union and Employer City Average Clearance Rates

Mas (2006) uses data on final offer arbitration rulings to test whether police offers "punish" communities after adverse bargaining outcomes

Reciprocity in the Labor Market: Summary

Don't expect much positive reciprocity from your employer, employees

- Workers, employers sometimes reward high effort
- Pattern not particularly robust
- Tendency to exert higher effort may wear off quickly

Negative reciprocity appears more reliable

Reciprocity: Theoretical Approaches

Rabin (1993): Fairness Equilibrium

Individual i's utility depends on:

- i's private payoff
- Beliefs about the "kindness" of j's chosen strategy
- Level of kindness (toward j) inherent in i's chosen strategy

Action space and belief structure:

- a_i is Player i's action (mixed strategy)
- b_j is Player i's belief about Player j's action
- c_i is Player i's belief about Player j's belief about i's action

Rabin (1993): Fairness Equilibrium

Kindness is defined in terms of how close Player i's action brings Player j to achieving her maximum possible (Pareto-efficient) payoff

$$f(a_i, b_j) = \frac{\pi_j(a_i, b_j) - \pi_j^{\text{fair}}(b_j)}{\pi_j^{\max}(b_j) - \pi_j^{\min}(b_j)}$$

where $\pi_j^{\text{fair}}(b_j) = \left[\pi_j^{\max}(b_j) - \pi_j^{\text{low}}(b_j)\right]/2$ is a neutral fair outcome

Utility takes the form:

$$U_i(a_i, b_j, c_i) = \pi_i(a_i, b_j) + \tilde{f}(b_j, c_i) \cdot [1 + f(a_i, b_j)]$$

where $\tilde{f}(b_j, c_i)$ measures how kind Player i believes Player j is attempting to be given her beliefs about i's strategy

Levine (1998): "Altruism and Spitefulness"

Levine (1998) proposes a model of altruism, spite, and reciprocity which incorporates beliefs about j's type into i's utility:

$$v_i = u_i + \frac{\alpha_i + \lambda\alpha_j}{1+\lambda} u_j$$

where $a_i, a_j \in [-1, 1]$ and $\lambda \geq 0$

The distribution of types is known, but Player j's type is unknown

- All games are now Bayesian games
- Player j's actions may reveal something about her type

Levine (1998) uses data from UGs to argue that $\lambda \neq 0$

Charness-Rabin (2002): a "Simple" Model

Player B's preferences can be represented by the utility function

$$u_b(\pi_a, \pi_b) = (\rho \cdot r + \sigma \cdot s + \theta \cdot q)\,\pi_a + (1 - \rho \cdot r - \sigma \cdot s - \theta \cdot q)\,\pi_b$$

with the model parameters defined as follows:

- π_a, π_b are payouts to a, b respectively
- $r = 1$ if $\pi_b \geq \pi_a$, zero otherwise
- $s = 1$ if $\pi_b < \pi_a$, zero otherwise
- $q = -1$ if Player A has misbehaved, $q = 0$ otherwise

Negative reciprocity pivots indifference curves, possibly to the point where Player B's utility is decreasing in the payoff to Player A

Cox et al (2007): a "Tractable Model"

Cox et al (2007) propose simple extension to CES utility:

$$u_i(\pi_i, \pi_j) = \left(\pi_i^\alpha + \theta(r)\pi_j^\alpha\right)/\alpha$$

The utility weight that Player i places on the payoff to Player j depends on i's "emotional state" — parameterized by $\theta(r)$

$$\theta(r_i) = \theta_0 + ar(x_j) + \varepsilon_i$$

and

$$r(x_j) = \frac{m_i(x_j) - m_i^{\text{fair}}}{\max_x m_i(x_j) - \min_x m_i(x_j)}$$

where $m_i(x_j)$ is the max payout to i if j chooses action x_j

$\Rightarrow$ $\theta(r)$ is strictly increasing in r