Interaction between selfish and conditionally altruistic players - a model suggested by public good experiments

(Work in progress)

Attila Ambrus

Harvard University, Department of Economics

PUBLIC GOOD EXPERIMENTS

subjects can decide how much money to contribute to public good and how much to retain / invest privately

retaining money dominates public contribution in terms of individual returns, but public contribution yields higher aggregate payoff

SIMPLE SPECIFICATION

N players, player i has endowment $z_i > 0$ and chooses public good contribution $t_i \in [0, z_i]$

payoff: $U_i(t) = z_i - t_i + a \sum_{j \in N} t_j$

where $a < 1$ but $Na > 1$

dominant strategy contribute 0 in 1 shot game, unique SPE is contributing 0 in finitely repeated game

EVIDENCE

instead in 1 shot game 40-60% contribution rate

if game is repeated, cooperation winds down

in last period roughly 75% of respondents contribute nothing, the rest very little

if game is restarted, contributions rise temporarily and wind down again (f.e. Andreani (88))

EXPLANATIONS

-learning story, initial mistakes

-reputation

-social norms adopted from the real world

have trouble explaining the patterns, especially the return to cooperation at the beginning of new rounds

the stability of the pattern suggests some kind of equilibrium story

ALTRUISM

has to be some form of altruism / fairness consideration, at least on the part of some players

Different specifications of altruism in the literature

Paternalistic/nonpaternalistic sympathy

Unconditional/conditional reciprocity

UNCONDITIONAL ALTRUISM/SPITE

A simple example of unconditional altruism/spite:

$$U_i = x_i + \sum_{j \neq i} \alpha_j x_j$$

CONDITIONAL ATTITUDES TOWARDS OTHERS

-well-being of the other players (care only about those worse off than me)

-care about one's relative standing

-equity

-combination of above

-intentions of others

RECIPROCICAL ALTRUISM

"If people do not think that others are doing their fair share, then their enthusiasm for sacrificing for others is greatly diminished" (Rabin)

PSYCHOLOGICAL GAMES

Geanakoplos, Pearce and Stachetti (89)

payoffs depend on both actions AND beliefs (expectations)

higher level beliefs can enter the payoff function (I care about what I expect other people to expect)

Psychological equilibrium: the expectations have to be correct and given these expectations players maximize

CONDITONAL KINDNESS IN PSYCHOLOGICAL GAME

Rabin: Incorporating fairness into game theory and economics (93)

First model to derive psychological game from underlying material game using the concept of conditional kindness

players care about material payoffs

plus want to treat those kind who (they think) treat them kind and treat those unkind who treat them unkind

Typically generates multiple equilibria (BoS game)

EXPERIMENTS SUGGEST HETEROGENEITY

Ledyard's observation (in Chapter 2, Handbook of experimental economics, ed. by Kagel and Roth) is that 50% of players pursue self-interest, 40% has fairness/altruism considerations besides self-interest but not unconditionally

Palfrey and Prisbey (93): 49% of players maximize individual payoffs

Saijo and Yamaguchi (92): at least half of the players close to behaving self-payoff maximizing

Fischbacher, Gachter & Ernst Fehr (01, Economics Letters): 50% of players are conditional cooperators

INNOVATION: EQUILIBRIUM BETWEEN SELFISH AND FAIR PLAYERS

"The interaction between fair and selfish individuals is key to the understanding of the observed behavior in strategic settings"

(Fehr&Schmidt: Theories of fairness and reciprocity - evidence and economic applications)

THE MODEL

2N players, players $1, ..., N$ selfish, players $N+1, ..., 2N$ conditionally reciprocal ("fair")

T periods

endowment 1 in each period

FAIR TYPES

their preferences depend on past contributions and current expected contribution (makes it a psychological game)

$\widetilde{t}^k_{-i} = \frac{1}{2N-1} \sum_{j \neq i} t^k_j$ the others' average contribution in period k

$\widetilde{e}^k_{-i}$ be i's expectation of $\widetilde{t}^k_{-i}$ at the beginning of k

$\widetilde{a}^k_{-i} = \frac{\sum_{j=1,\ldots,k-1} \widetilde{t}^j_{-i} + \widetilde{e}^k_{-i}}{k}$ expected per-period contribution of others in the first k periods

$\widetilde{a}^1_{-i} = \widetilde{e}^1_{-i}$

UTILITY FUNCTION FOR FAIR TYPES

Then i's period k payoff is $U_i^k(\tilde{a}_{-i}^k, t_i^k) = 1 - t_i^k + at_i^k + f(\tilde{a}_{-i}^k - t_i^k)$

where $f' \leq 0$

material payoff plus psychological cost of contributing little compared to others' expected per period contribution

Total payoff is sum of per period payoffs

CONCRETE SPECIFICATION

$$f(x) = \begin{cases} 0 \text{ if } x > 0 \\ bx \text{ if } x \leq 0 \end{cases}$$

where $b > 1 - a$

Then the contribution that maximizes i's per period payoff is $t_i^{k*}(\widetilde{a}_{-i}^k) = \widetilde{a}_{-i}^k$

SOLVING THE 2 PLAYER, 2 PERIOD MODEL

S -selfish player, F -fair player

Last period: S contributes 0, F contributes $\frac{t_1^1}{2}$

In first period S cannot influence what F does this period, but can influence next period

Then $t_1^1 = 1$ if $(a + 1/2)t_1^1 > 1$

$t_1^1 = 0$ if $(a + 1/2)t_1^1 < 1$

t_1^1 arbitrary if $(a + 1/2)t_1^1 = 1$

With the exception of the last case unique equilibrium

In first case total contribution is 2 in the first period and 1/2 in the second

2 PLAYERS, T PERIODS

same logic establishes that S at any period k considers only a contribution's effect on future periods

contribution d in period k increases P1's utility in later periods by $\frac{d}{k+1} + ... + \frac{d}{T}$

Then P1 contributes 1 at periods $1, ..., K$ and 0 afterwards

where K is the smallest integer not larger than K for which $\frac{c}{K+1} + ... + \frac{c}{T} + a < 1$

Then P2 is a psychological equilibrium contributes 1 at periods $1, ..., K$ and then $\frac{K}{K+1}, ..., \frac{K}{T}$

Again psychological equilibrium is generically unique

GENERAL RESULTS

If f is such that t_i^{k*} is linear in $\widetilde{a}_{-i}^k$ then for every N and T

-psychological equilibrium generically unique

-total contribution is (weakly) decreasing

-selfish players first contribute then at some point they stop, in particular in last period contribute 0

The latter two can be established for much more general f (work in progress)

CONTRASTING PREDICTIONS WITH EXPERIMENTAL RESULTS

-Increase in a increases total contribution consistent with experimental evidence

Isaac et al (84), Kim & Walker, Isaac & Walker (88)

-Relatively high stakes don't change the patterns very much

Hoffman et al, Slonim and Roth

INCREASING THE NUMBER OF PLAYERS

Increasing N decreases the value of contributing for selfish types

In above example contribution d in period k increases S's utility in later periods by $\frac{d}{(2N-1)(k+1)}+...+\frac{d}{(2N-1)T}$

which decreases in N, so selfish types stop contributing earlier

But nonselfish players face a higher fraction of nonselfish others who continue contributing (if N=4 then 1/3, if N=10, then 4/9). Their contribution decreases at a slower rate

Bagnoli & McKee (91): increasing N effects cooperation negatively, especially in early perods

MULTIPLE EQUILIBRIA, COMMUNICATION

There can be multiple equilibria if $t_i^{k*} / \frac{\sum_{j=1,\dots,k-1} \widetilde{t}^j_{-i} + \widetilde{e}^k_{-i}}{k}$ decreases in $\frac{\sum_{j=1,\dots,k-1} \widetilde{t}^j_{-i} + \widetilde{e}^k_{-i}}{k}$ (for example if $t_i^{k*} = 1$ if $\frac{\sum_{j=1,\dots,k-1} \widetilde{t}^j_{-i} + \widetilde{e}^k_{-i}}{k} > 1/2$ but $t_i^{k*} = 0$ otherwise; relatively high average contribution is rewarded, and positive but relatively low contribution is not)

In these cases multiple equilibria and room for coordination (which is shown to have a positive effect on total contributions)