Endogenous Institution Formation in Public Goods Games

Based on Kosfeld, Okada, and Riedl (2009)

Mingzhao Dai

12.11.2024

Introduction

- Public goods games + endogenous institution formation.
- Standard preferences vs. social preferences.
- Theoretical model + experimental design.

Model Setup

- n-player public goods game.
- Each player has an endowment w.
- Contribution to public good: $g_i \leq w$.
- Payoff for player i:

$$\pi_i(g_1, \ldots, g_n) = w - g_i + a \sum_{j=1}^{n} g_j$$

  where $a < 1$ is the marginal per capita return (MPCR).
- Given $a < 1$, the NE is for all players to free-ride ($g_i = 0,\ \pi_i = w$).
- Additionally assume $na > 1$, so that (w,...,w) is better than (0,...,0) and is also welfare-maximizing ($g_i = w,\ \pi_i = naw$).

(Endogenous) Institution Formation

1. **Participation Stage**: Players decide whether to join a sanctioning institution.
2. **Implementation Stage**: Participants decide if the organization will be implemented.
   (unanimous consent: if anyone does not agree, the institution cannot be implemented)
3. **Contribution Stage**: Players contribute after implementation; institution members face sanctions if they under-contribute.

Player's Final Payoff and Sanction Mechanism

Final Payoff Formula:

$$u_i = \begin{cases} w - g_i + a\sum_{j=1}^{n} g_j - \frac{c}{s} - p(g_i), & \text{if } i \in S \text{ (member)} \\ w - g_i + a\sum_{j=1}^{n} g_j, & \text{if } i \notin S \text{ (non-member)} \end{cases}$$

- c: Total cost of maintaining the institution.
- $s = |S|$: Number of players in the institution.
- $p(g_i)$: Penalty applied to member i if they under-contribute.

Sanction Mechanism (for only members):

$$p(g_i) = \begin{cases} w - g_i, & \text{if } g_i < w \quad \rightarrow \textit{goes to nowhere} \\ 0, & \text{if } g_i = w \end{cases}$$

- Members always contribute all endowments ($g_i = w,\ i \in S$).

Standard Preferences

- 3rd stage: Players contribute w if institution is implemented.
- 2nd stage: Players implement the institution if

$$asw - \frac{c}{s} > w$$

- Subgame perfect Nash equilibria (SPNE):
  - **Organizational Equilibrium**: Institution is formed with

$$s \geq s^*,$$

    where threshold s^*: minimum group size needed for institution[1].
    That means, the institution is formed only if there are sufficient members.
  - **Status Quo Equilibrium**: No institution forms if not sufficient players join ($s < s^*$).

[1] Under certain harmless conditions, $2 \leq s^* \leq n$.

Standard Preferences: Refinement

- SPNE tells us that there exists multiple equilibria (e.g. any $s \geq s^*$ forms an organizational equilibrium).
- To refine the equilibria set, we introduce the definition of **strict** SPNE: best response at every stage is unique (no deviation).
- Unique strict SPNE:

$$s = s^*$$

.

- Intuition: players will not join the institution (at the 1st stage) if they know others will form an institution (conditionally free ride).
- Strict SPNE is the most plausible organizational equilibrium.

Social Preferences: Inequity Aversion

Utility Function for player i (Based on Fehr and Schmidt, 1999):

$$U_i = u_i - \alpha_i \cdot \frac{1}{n-1} \sum_{j \neq i} \max(\pi_j - \pi_i, 0) - \beta_i \cdot \frac{1}{n-1} \sum_{j \neq i} \max(\pi_i - \pi_j, 0)$$

- u_i: Player i's material payoff.
- α_i: Degree of **disadvantageous** inequity aversion.
- β_i: Degree of **advantageous** inequity aversion.

Incentive to implement now changes:

$$asw - \frac{c}{s} - \frac{\alpha_i}{n-1}(n-s)(w + \frac{c}{s}) > w$$

- New threshold $s^+_{(i)} > s^*$: larger group size needed.
- $s^+_{(i)}$ increasing in α_i: stronger inequity aversion, larger size.
- Any $s^+_{(i)} \leq s \leq n$ forms an organizational equilibrium.

Social Preferences: Grand Organization

Again, we do the refinement for the equilibria set.

Results: **Grand organization** ($s = n$) is the most plausible organizational equilibrium. Under the following conditions, it is the strict SPNE, and may be even unique: [2]

- (Strictness) At least for 2 players, $\alpha_i > \tilde{\alpha}$.
- (Uniqueness) At least for $n - 1$ players, $\alpha_i > \tilde{\alpha}$.

Note that $\tilde{\alpha}$ is increasing in a.

Illustration: Stronger **inequity aversion** is more likely to lead to the formation of a grand organization where everyone participates.

[2] Under the condition that β_i is relatively small. Similar propositions hold for larger β_i with modified conditions. $\tilde{\alpha}$ is a constant given $\{n, a, w, c\}$.

Experimental Settings

- Subjects play a 20-round, four-player public goods game (same group).
- Each player has an initial endowment of 20 points.
- The procedure follows the model, except for the contribution stage (i.e. fixed $g_i = w$ for members $i \in S$).
- Introducing the "belief" of the group size.
- Experiment includes two experimental treatments:
  1. **IF40** - MPCR $a = 0.4 \quad \Rightarrow s^* = 3$.
  2. **IF65** - MPCR $a = 0.65 \quad \Rightarrow s^* = 2$.
- Control treatments (**PG40** and **PG65**) included for comparison, without institution formation options.

Experimental Results

- Majority of organizations formed were "grand organizations" where all players joined.

TABLE 1—INITIATED AND IMPLEMENTED ORGANIZATIONS

	Treatment			
	IF40		IF65	
	Number	Percentage	Number	Percentage
Initiated organizations	220	100	216	98
Implemented organizations				
Total	95	43	132	61
One member	0	0	5	4
Two members	1	1	15	11
Three members	15	16	22	17
Four members	79	83	90	68

Figure: Distribution of organization sizes in IF40 and IF65 treatments.

Comments: 1. Players rarely formed organizations under the size of s^*. 2. Outcomes approached n instead of s^*, meaning the social preferences are playing an important role.

Experimental Results

- Players learned to form grand organizations over time.

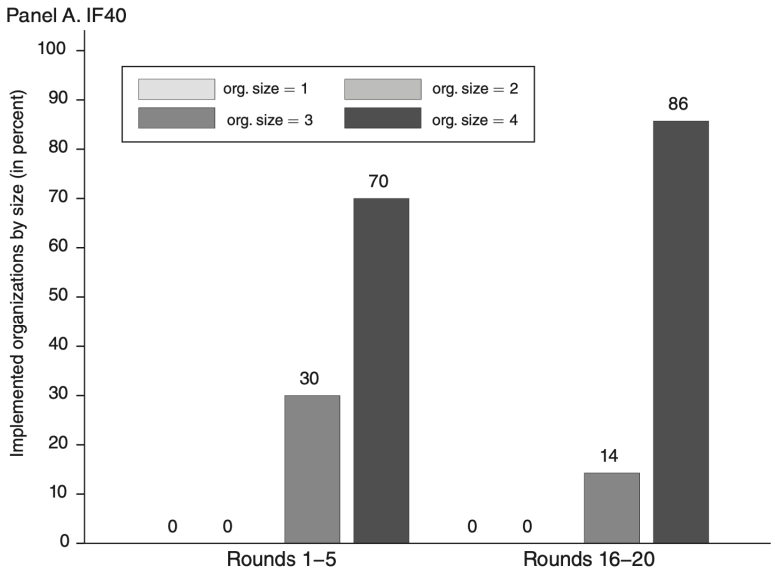

Figure: Distribution of organization sizes in IF40 over time.

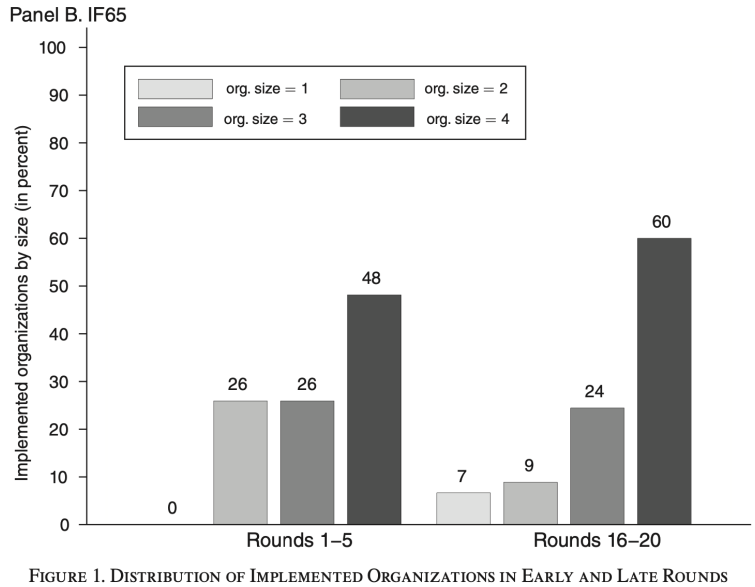

FIGURE 1. DISTRIBUTION OF IMPLEMENTED ORGANIZATIONS IN EARLY AND LATE ROUNDS

Figure: Distribution of organization sizes in IF65 over time.

Experimental Results

- Grand organizations are more stable (higher implementation rate).

TABLE 2—BELIEFS AND RATE OF IMPLEMENTATION

	Treatment							
	IF40				IF65			
	Number of participants				Number of participants			
Implementation rate	1	2	3	4	1	2	3	4
All rounds	0.00	2.94	23.08	69.30	27.78	37.50	37.29	90.91
Observations	7	34	65	114	18	40	59	99

Figure: Implementation rate of different sizes in IF40 and IF65 treatments.

Comments: Higher a $\Rightarrow$ higher α_i $\Rightarrow$ Uniqueness condition more likely to be held $\Rightarrow$ more stable.

Experimental Results

- Possibility of institution formation significantly increased contributions and total welfare.

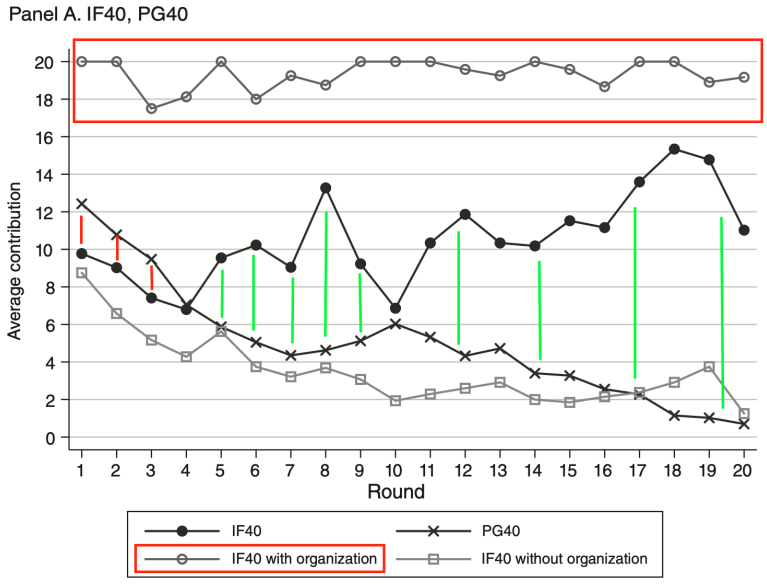

Figure: Average contributions in IF40 over time.

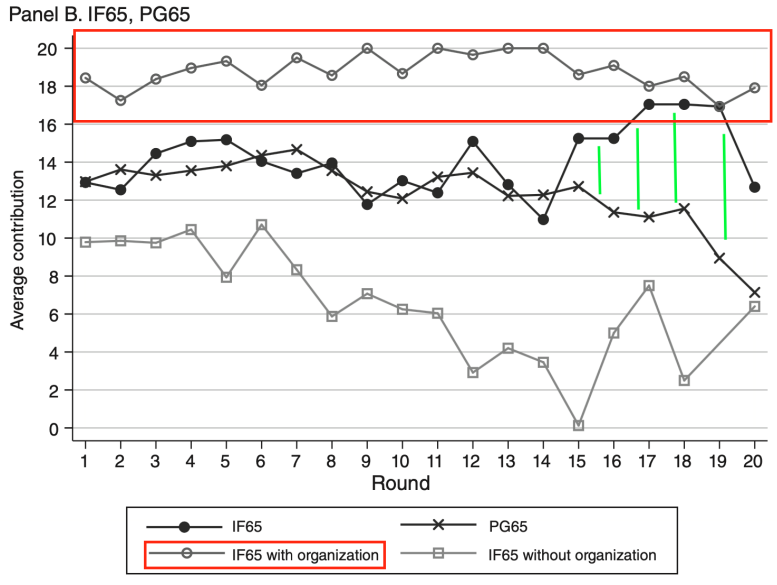

Figure: Average contributions in IF65 over time.

Comments: Interesting fact: Total contributions between IF65 and PG65 are similar in the first 14 rounds. (Institution formation failure has a negative effect on voluntary contributions.)

Critical Reviews

- **Strengths**:
  - **Simple and powerful**: using classic public goods game model and also classic Fehr and Schmidt (1999) inequity aversion model to explain new intuitions.
  - **Elegant**: combining both theoretical and experimental studies, and more elegantly, experimental design effectively tests the model's predictions.

Critical Reviews

- **Weaknesses**: Results driven by the game structure design.
  - **Multi-stage game** is bad:
  - **"Hold-up" problems may arise**:

    Players who naively participate at the 1st stage are forced to implement at the 2nd stage due to the requirement for "unanimous consent". The rest of the players cannot form the institution by themselves.
  - **"Implementation stage" is just designed for the sake of the model, to allow players to see the size** s.

    If we can already see the size of the institution, it's of course easy to make decisions. We don't need the model to tell the "insights".
  - **Incentives for participation (1st stage) unclear**:

    With standard preferences, s^* players who will participate are picked up by an invisible hand. With social preferences, players should also not observe their expected payoff at the 1st stage. Then why they participate? The experiment uses "belief" to solve this incentive problem.
- In my mind, the endogenous institution formation should be a "social contracting" problem, which is more like a one-shot game. Players should have beliefs and incentive compatible constraints.

Links to the course

- *"Fehr-Schmidt fits public good game without punishment reasonably well"*
- *"Smaller groups are more likely to reach a cooperative equilibrium"*
- *"Andreoni (AER, 1995) tries to separate kindness and confusion ... in VCM"*, *"Initial high contributions could be a result of misunderstanding the game. Decline of contributions would then result from learning"*

Thank you!