

What are the equilibria in public-good experiments?

Irenaeus Wolff

Thurgau Institute of Economics (TWI) / University of Konstanz, Hauptstrasse 90, 8280 Kreuzlingen, Switzerland

HIGHLIGHTS

- I determine the Nash-equilibria resulting from participants' elicited preferences.
- Multiple equilibria are relatively frequent even in a standard 3-player setting.
- Multiple equilibria are more frequent than Fehr & Schmidt's (1999) model predicts.
- Hence, omnilateral defection at the end of repeated public-good games is surprising.

ABSTRACT

Most social-preference models have been tailored to yield only a full-defection equilibrium in one-shot linear public-good situations. This paper determines the Nash-equilibrium sets that result from experiment participants' elicited preferences. The data show that multiple equilibria are relatively frequent even in a standard three-player setting. In this perspective, the common finding of close-to-omnilateral defection at the end of repeated public-good games is surprising and raises the question of why the dynamics of play seem to select this equilibrium out of the existing equilibria.

1. Introduction

In many social-game protocols, ranging from gift-exchange over ultimatum bargaining to public-good situations, human behaviour differs substantially from the Nash-equilibrium that results if we assume that players care only about their own monetary payoff. To resolve this discrepancy, numerous models of social preferences have been proposed.¹ These models have been tailored to fit stylised facts from the laboratory. To take a prominent example, Fehr and Schmidt (1999) took it as support for their model that it singles out full-defection as the virtually-unique equilibrium in the typical public-good setting, given most experiment participants cease to contribute in the final round(s) of repeated-play experiments. Yet, no study has documented empirically how often omnilateral defection is an equilibrium given participants' prefer-

ences, and how often there are additional, positive-contributions equilibria. This is what the present paper does.

This paper elicits preferences under three different sets of parameters of a linear public-good protocol, and documents the Nash equilibria of all potential within-treatment matchings of participants, which I call *revealed-preference Nash-equilibria* (RPNE). Thereby, I provide empirical evidence of the strategic environment induced by different public-good protocols that can be contrasted with theory.²

2. Preferences and equilibria

In line with preceding studies, I use a reduced-form approach to preferences: I look at conditional-contribution preferences,

E-mail address: wolff@twi-kreuzlingen.ch.

¹ E.g., Rabin (1993), Fehr and Schmidt (1999), or Levine (1998).

² Following Weibull (2004), I use the term *public-good protocol* to denote a situation in which the *material consequences* have a public-good structure. Whether this translates into a *public-good game* then depends on participants' preferences.

Table 1
Overview of the treatments.

Treatment	n	E	m	MPCR
3P.5	3	20	1.5	0.5
2P.75	2	20	1.5	0.75
2P.67	2	15	4/3	2/3

that is, how much participants are willing to contribute to the public good depending on others' contributions. Fischbacher et al. (2012) show that this approach is behaviourally valid in the sense that contributions in a simultaneous public-good experiment can be explained by participants' elicited conditional-contribution preferences in conjunction with their beliefs.³ A pure-strategy revealed-preference Nash-equilibrium (RPNE) of the simultaneous game then is a contribution profile in which each player chooses a contribution in line with her conditional-contribution preferences given her belief about the other players' contributions, and beliefs match the respective other players' contributions.

3. Experimental design

I use data from three one-shot treatments designed to elicit conditional-contribution preferences by the method introduced in Fischbacher et al. (2001) and refined in Cheung (2013).⁴ Table 1 lists the parameters of the three treatments, where n is group size, E is participants' endowment in Euros, m is the public-good multiplier, and the resulting marginal per-capita return is abbreviated to MPCR. I denote treatments using the pattern $nP(\text{erson})+MPCR$: for example, the 3P.5-treatment is a 3-person protocol with an MPCR of 0.5. In all treatments, choices were restricted to six contribution levels. To stick to six levels while keeping profit calculations simple for participants also in the 2P.67-treatment, E had to be adjusted along with m .

I restricted contributions to six levels because of the importance to elicit the full conditional-contribution vector in the three-player treatment.⁵ It is essential to elicit responses to all contribution combinations because the players' response to contributions of, e.g., (8, 8) may be very different from their response to (0, 16). For example, a participant with a Fehr–Schmidt utility function would choose 0 in response to (0, 16), but 8 in response to (8, 8) as long as $\beta > 0.5$.

Prior studies have minimised participants' confusion about the situation by looking at behaviour at the end of repeated-game experiments. I use a different approach, inviting only experienced participants.⁶ Table 2 shows an overview of the sessions by treatment. All sessions were conducted at the University of Konstanz' LakeLab between January 2012 and January 2016, using z-Tree (Fischbacher, 2007) and ORSEE (Greiner, 2015). No participant participated more than once.

³ See also my working paper Wolff (2015) that is partially based on the same data as this letter.

⁴ In contrast to these papers, I presented first-mover contribution(combination)s one by one, in an order that was randomised individually for each player. All three treatments were part of sessions with multiple parts and random rematching between parts. Participants were paid for one randomly-chosen part only, and parts were explained only as soon as they began. There was no feedback on earlier parts before the preference-elicitation treatments, so that I focus on these treatments here. For a detailed discussion of the full experimental setup of what I will call the 3P.5- and the 2P.75-treatments (including the full set of instructions for 3P.5), cf. Wolff (2015). The session setup of the 2P.67-treatment followed a very similar design, the most important difference being that the 2P.67-treatment was the second rather than the fifth part in the session.

⁵ See Cheung (2013).

⁶ Participants in the experiment had participated in at least one public-good experiment and at least four additional other experiments, with no upper limits.

Table 2
Overview of the sessions by treatment.

	3P.5	2P.75	2P.67
Number of sessions	10	3	4
Participants	236	76	82

4. Results

From the literature, we know there is considerable heterogeneity in conditional-contribution preferences. Table 3 shows the distribution of preference-types introduced by Fischbacher et al. (2001, for the classification procedure, see the Online Appendix), alongside the corresponding distribution in each of the treatments of this study. Except for an unusually high fraction of 20% unclassifiables in 2P.67, the distributions are close to what we would expect: because cooperation gets cheaper, a higher MPCR leads to more conditional cooperation and less defection for a fixed group size (2P.67 vs 2P.75), while increasing the group size with (almost) constant multiplier m (2P.75 vs 3P.5 vs Fischbacher et al.'s '4P.4'), has the fraction of conditional cooperators steadily decline as the fraction of defectors increases.⁷

Using the elicited conditional-contribution vectors, I calculate the pure-strategy RPNE sets of all potential matchings within each treatment as detailed in Section 2. I then classify the RPNE set for each of these hypothetical groups according to the cardinality of the RPNE set and according to whether they include full-defection/low-contributions equilibria and high-contributions equilibria. The description of the chosen RPNE set classes and their prevalence in a perfectly-randomised sample are given in Table 4, along with the predicted distribution for the calibrated model of Fehr and Schmidt (1999), as an exemplary benchmark.

Four RPNE-set classes account for 86%–93% of all RPNE sets to be expected: (i) a unique, full-defection RPNE, (ii) a unique positive-contributions RPNE (with average contribution levels of 40%–45% irrespective of the treatment), (iii) multiple RPNE that range from full-defection to high contributions, and (iv) multiple RPNE with average contributions of always less than half the endowment including a full-defection RPNE. Note that the importance of the type-(ii) RPNE class differs widely between treatments. In particular, it seems to play a substantial role only in the 2P.67-treatment were by some chance, we had an unusually-high percentage of unclassifiable participants. What is important here is that there is a surprisingly high prevalence of multiple-RPNE sets in a well-mixed population for all treatments. For example, the prevalence of multiple-RPNE sets clearly exceeds the predicted frequency on the basis of the Fehr–Schmidt model (see the final row in Table 4). On the other hand, this model does seem to capture the comparative statics between treatments for the two RPNE-set classes that are most prevalent overall, type-(i) and (iii). Note also that—as predicted by all commonly-used social-preference models including Fehr and Schmidt (1999)—the vast majority of all possible matches lead to a RPNE set that includes full-defection.

⁷ The latter comparison mirrors differences in contribution levels, e.g., between the "LOW_8" and "HIGH_3" treatments in Nosenzo et al. (2015). However, a regression of participants' reactions to full contributions by their peers lends support only to the MPCR effect ($p = 0.077$) but not to the group-size effect ($p = 0.896$; note that we could not include Fischbacher et al.'s data). In addition, economics students contribute less in response to others' full contributions ($p = 0.001$), females (insignificantly) contribute more ($p = 0.148$). Participants' intelligence (as measured by their average A-levels grade), their self-reported religiousness, and whether they work or not do not affect their reactions to full contributions.

Table 3
Distribution of player types.

Treatment	Percentage of...			
	Conditional cooperators	Defectors	Triangle cooperators	Others
3P.5	60	23	11	6
2P.75	76	16	4	4
2P.67	48	21	12	20
Fischbacher et al. (2001); '4P.4'	50	30	14	7

Table 4
Classification and expected distribution (in percent) of RPNE-set types. Predictions of the calibrated Fehr–Schmidt model are added in parentheses.

RPNE-set type	Description	3P.5	2P.75	2P.67
\emptyset	No pure-strategy RPNE	0.1 (0.0)	1.7 (0.0)	4.2 (0.0)
$\{(0, 0, 0)\}$	Unique RPNE characterised by full defection by all group members	60.1 (93.6)	29.8 (51.0)	38.1 (84.0)
$\{(x, y, z)\}$	Unique RPNE where at least one group member's contribution is strictly positive	2.1 (0.0)	11.5 (0.0)	26.4 (0.0)
FULLD-LIMITED	A full-defection RPNE and at least one additional RPNE; the RPNE with the highest average contributions has an average contribution of less than half the endowment	9.6 (0.0)	6.7 (0.0)	5.7 (0.0)
FULLD-INTERMED	A full-defection RPNE and at least one additional RPNE; the RPNE with the highest average contributions has average contributions of between 50% and 80% of the endowment	5.1 (0.0)	3.0 (0.0)	3.4 (0.0)
FULLD-HIGHC	A full-defection RPNE and at least one additional RPNE in which players contribute at least 80% of their full endowment on average	21.4 (6.4)	38.1 (49.0)	17.4 (16.0)
LOWC-HIGHC	A high-contributions RPNE (s.a.) and at least one additional RPNE with average contributions of at most 20% of endowment	0.5 (0.0)	4.9 (0.0)	0.9 (0.0)
ONLYHIGH	At least two RPNE, in all of which average contributions are higher than half the endowment	0.4 (0.0)	0.0 (0.0)	0.7 (0.0)
OTHERS	Multiple-RPNE sets that do not fit any of the above categories (88%/51%/41% of these sets include full-defection)	0.7 (0.0)	4.3 (0.0)	3.2 (0.0)
MULTIPLE	Cumulated percentage of all multiple-RPNE sets	37.7 (6.4)	57.0 (49.0)	31.3 (16.0)

5. Discussion

In this paper, I documented the distribution of equilibrium-set classes that typical participants would face in the laboratory when presented with linear public-good protocols. The finding that the prevalence of multiple (high-cooperation) equilibria may be higher than commonly expected underlines the necessity of conducting this type of exercise also for other situations in which social preferences are thought to be important. Taking into account the prevalence of multiple equilibria, the ubiquitousness of close-to-omnilateral defection at the end of repeated public-good protocols is surprising and re-opens the question of why the dynamics of play seem to select this equilibrium out of the existing equilibria.

Acknowledgements

I am particularly grateful to Urs Fischbacher for many helpful discussions and for the idea for the algorithm to calculate the equilibria in virtually no time. Furthermore, I would like to thank the lively research group at the Thurgau Institute of Economics (TWI), Alexander K. Wagner, Simon Gächter, and participants of the ESA European Meeting 2015, the 15th International Conference on Social Dilemmas, and seminar audiences in Vienna, Innsbruck, and Nottingham for fruitful discussions. I thank the Swiss National Foundation for funding under grant 100018_152788/1.

References

- Cheung, S.L., 2013. New insights into conditional cooperation and punishment from a strategy method experiment. *Exp. Econ.* 1–25.
- Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114 (3), 817–868.
- Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. *Exp. Econ.* 10 (2), 171–178.
- Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Econom. Lett.* 71 (3), 397–404.
- Fischbacher, U., Gächter, S., Quercia, S., 2012. The behavioral validity of the strategy method in public good experiments. *J. Econ. Psychol.* 33 (4), 897–913.
- Greiner, B., 2015. An online recruitment system for economic experiments. *J. Econ. Sci. Assoc.* 1 (1), 114–125.
- Levine, D.K., 1998. Modeling altruism and spitefulness in experiments. *Rev. Econ. Dyn.* 1 (3), 593–622.
- Nosenzo, D., Quercia, S., Sefton, M., 2015. Cooperation in small groups: the effect of group size. *Exp. Econ.* 18 (1), 4–14.
- Rabin, M., 1993. Incorporating fairness into game theory and economics. *Am. Econ. Rev.* 83 (5), 1281–1302.
- Weibull, J.W., 2004. Testing game theory. In: Huck, S. (Ed.), *Advances in Understanding Strategic Behaviour: Game Theory, Experiments and Bounded Rationality. Essays in Honour of Werner Güth*. Palgrave Macmillan, pp. 85–104.
- Wolff, I., 2015. When best-replies are not in equilibrium: Understanding cooperative behaviour. Research Paper Series 97, Thurgau Institute of Economics.