Not efficient but payoff dominant
Experimental investigations of equilibrium play in binary threshold public good games

Friedel Bolle

Jörg Spiller

European University Viadrina Frankfurt (Oder)

Department of Business Administration and Economics

Discussion Paper No. 379

February 2016

ISSN 1860 0921

Not efficient but payoff dominant

Experimental investigations of equilibrium play in binary threshold public good games

Friedel Bolle, Jörg Spiller

In Binary Threshold Public Good (BTPG) games players contribute or not to the production of a public good which is produced if and only if there are "enough" contributors. There is a plethora of equilibria in BTPG games. We experimentally test general theoretical attributes of equilibria and proposals for equilibrium selection. As theory predicts, if the cost/benefit ratio is the same, then subjects play (almost) the same mixture of strategies and, after switching from a positive to a negative frame, the theoretically expected "mirrored" behavior can be observed, i.e. contrary to most linear Public Good experiments we do not find a framing effect. A finite mixture model successfully (i.e. without rejection in a chi-square test) describes behavior in all eight experimental games (same parameters for four thresholds and positive/negative frame). The Harsanyi-Selten theory of equilibrium selection is moderately supported. Efficiency as an equilibrium selection device and also risk dominance are clearly rejected.

JEL-Classification: C72, D72, H41

Keywords: Binary threshold Public Goods, framing, equilibrium selection, payoff dominance, risk dominance, efficiency, experiment

Highlights

- a public good is produced if a sufficient number of players support it
- players with the same cost/benefit ratio contribute with the same frequency
- negative vs. positive costs and benefits do not cause a framing effect
- payoff dominance and the Harsanyi-Selten equilibrium selection are supported
- risk dominance and efficient equilibrium play are rejected

1. Introduction

In order to substitute the empty toner cartridge of a publicly used printer only one volunteer is needed who is ready to bear the costs in terms of time lost and dirty hands. This is an example of the Volunteer's Dilemma, first analyzed by Diekmann (1985). It requires all members of a cartel to keep their contract secret. With plausible assumptions about the profitability of the cartel and incentives for Whistle Blowers, this is an example of the Stag Hunt game, first described by Rousseau (1997). These are the extreme cases of Binary Threshold Public Good (BTPG) games where players have a dichotomous choice of either contributing or not. More general thresholds require fixed contributions with costs $c_i > 0$ by at least k of n members of a group in order to produce a public good with profits $G_i > c_i, i = 1, ..., n$. This structure is completely different from a linear public good game with binary contributions which has a unique equilibrium (no one contributes) while a BTPG game has a plethora of pure and mixed strategy equilibria. If $G_i < c_i < 0$, then it is individually profitable to contribute but players provide a "public bad" when contributions surpass the threshold.

Voting in parliaments and in committees as well as shareholder voting is an important example for BTPG games[1] (see Bolle, 2015a). In many other examples a team with minimal requirements concerning the number and perhaps also the complementary qualifications 1 of the members is necessary to launch a project or solve a problem for the best of their community . We report here about BTPG experiments with essentially symmetric players. Players have different positive costs c_i and benefits G_i but their cost/benefit ratio is the same. In a second treatment, costs and benefits are negative but absolutely the same as in the first treatment. There are n=4 members of the experimental groups. In both treatments, the thresholds are described by a minimum number $k = 1, 2, 3$ or 4 of contributors.

Our goal is to understand behavior in these games, mainly by testing game theoretic predictions. We first test two theoretically derived attributes: (i) Players with the same cost/benefit ratio behave equally. (ii) In a negative frame ($G_i < c_i < 0$), we observe the "mirrored" behavior of that in the positive frame ($0 < c_i < G_i$). Mirroring means, that contributing and non-contributing are exchanged and the threshold is changed from k to $n-k+1$. We find that these theoretical predictions are (almost) fulfilled. Note that it is a bit surprising that these regularities are confirmed because framing effects are regularly reported in economic experiments. In linear Public Good experiments, it has always been found that the negative frame (linear static Common Pool experiments) evokes significantly less cooperation than the positive frame[2] (Andreoni, 1995; Willinger and Ziegelmeyer, 1999; Park, 2000; Dufwenberg et al., 2011). This result is confirmed in the only BTPG experiment with a positive and a negative frame (Sonnemans et al., 1998, see below).

Our next goal is rather general, namely (iii) to present a description of behavior as (mainly)

[1] "Complementary qualifications" require a generalization of the threshold definition. In Bolle (2015a) the threshold is described by sufficient subsets of players as in cooperative games with binary characteristic functions (called simple cooperative games or voting games).

[2] If we investigate dynamic games with incremental contributions the giving to a public good and withdrawing contributions from an initial state of full contributions are no longer frames but define different games. Breitmoser et al. (2015) show the theoretical as well as behavioral Pareto-superiority of the taking game.

equilibrium play with stochastic deviations. The results of the eight games ($k = 1, 2, 3, 4$ and positive/negative frame) are explained by the same mixture of behavioral modes where the equilibrium selected by the Harsany-Selten theory of equilibrium selection, called HS, takes a share of 49%. The HS equilibrium is characterized as symmetric and payoff-dominant among the symmetric equilibria. Competing modes of play are described in Section 5. Neither risk dominance nor efficiency seems to play an important role. In a χ^2-test, the finite mixture model shows a "perfect fit", i.e. it cannot be rejected ($p = 0.11$). A supplementary regression, however, shows that the static model has missed an important dynamic effect: Each game is repeated eight times in a stranger design and when a player was "pivotal" in the previous round then his contribution probability increases in the positive and decreases in the negative frame. Pivotality of player i means that $k-1$ of i's co-players had contributed. Related findings in the literature stem from experiments with sequential contributions (see below) where pivotality means that, without a player's contribution, k cannot be reached. The difference is that, with sequential contributions, players know when they are pivotal while, with several rounds of the static game, players experience to be pivotal and seem to believe that, even in a stranger design, i.e. with new co-players, there is an increased probability of being pivotal also in the next round.

To the best of our knowledge, (i) has never been checked in the literature and (ii) only in one paper (in a different experiment and with different results, see below). In neither is there an attempt to describe behavior by a general approach as in (iii).

Experimental studies of the Volunteer's Dilemma with equal cost/benefit ratios are (Diekmann, 1985; Franzen, 1995; Goeree et al., 2005). An important result is that, contrary to the theoretical prediction, the probability of success does not decrease with group size. Diekmann (1993) rejects the theoretical prediction that players with higher cost/benefit ratios use mixed strategies with higher mixture probabilities. In Public Good experiments with a punishment option (first investigated by Fehr and Gächter, 2002), punishment can be given the structure of the Volunteer's Dilemma, i.e. a punisher causes a predetermined loss for the punished player and further punishers do not increase the loss. Przepiorka and Diekmann (2013) and Diekmann and Przepiorka (2015) investigate such situations with different costs of the players and find an (incomplete) coordination on the lowest cost player as a volunteer, i.e. the asymmetric efficient equilibrium is supported. We ask whether these results can be replicated and extended to higher thresholds: We will test the predictive value of an equilibrium selection hypothesis E with the selection of the efficient[3] (group income maximizing) equilibrium. In a finite mixture model, however, we estimate a zero share for this mode of behavior.

Experimental studies of the Stag Hunt game are mostly based on 2x2 games and are concentrated on the question which of the two pure strategy equilibria "no one contributes" (mostly selected by risk dominance) and "all contribute" (selected by payoff dominance) is played. Van Huyck et al. (1990) and Rydval and Ortmann (2005) find tendencies towards risk dominance; tendencies towards payoff dominance are found by Battalio et al. (2001), provided the "optimization premium" is high enough, and, in an experiment with chimpanzees, by Bullinger

[3]Indications for efficiency concerned behavior in Dictator experiments with choices from a finite set of income distributions are Kritikos and Bolle (2001) and Engelmann and Strobel (2004).

et al. (2011). Whiteman and Scholz (2010) find a positive influence of social capital. Al-Ubaydli et al. (2013) find that cognitive ability and risk aversion have no impact on successful coordination while patience does. Büyükboyacı (2014) shows that information about the risk attitude of others changes behavior which is, however, not affected by one's own risk attitude. Our experiments with four-player games are difficult to compare with all of these two-player games. Our results strongly support payoff dominance in the Stag Hunt game (74% - 93% contributions). Feltovich et al. (2013) investigate the influence of group size (2 to 7 players) and communication on contributions. Without communication, contributions are independent of group size (between 37% and 45%).

Experiments with intermediate thresholds require contributions from two of three players up to six of ten. In all investigations only one or two different thresholds are considered. Goren et al. (2003) investigate a BTPG game with five players with different weights $(5, 10, 15, 20, 25)$ and a threshold which requires the sum of weights to be at least 30. With the exception of Palfrey and Rosenthal (1991) all experiments are with complete information about monetary payoffs. Van de Kragt et al. (1983) and Palfrey and Rosenthal (1991) emphasize the importance of communication for successful coordination. Dawes et al. (1986) and Rose et al. (2002) investigate the (positive) influence of refunds of insufficient contributions and Dawes et al. (1986) also the punishment of successful free riding. Goren et al. (2003) find that the sequential-moves game leads to more efficient outcomes than the simultaneous-moves game and Erev and Rapoport (1990) show that, in addition, the information provided to the players in the sequential game matters. Erev and Rapoport (1990), Chen et al. (1996), and McEvoy (2010) find that, in sequential decisions, the pivotality (criticality) of players increases the contribution frequency. Bartling et al. (2015) find that pivotality increases responsibility attribution. Sonnemans et al. (1998) is, to the best of our knowledge, the only BTPG experiment where players contribute under a positive and under a negative frame. In the course of repetitions of games they find, in the negative frame, a trend towards less cooperation while we find a tendency towards HS, i.e., more cooperation. The main difference between the experiments is that Sonnemans et al. (1998) investigate 20 rounds with a partner design while we have eight rounds with a stranger design (i.e. a new random selection of groups in every round). In addition, their experiment has a threshold of $k = 3$ of $n = 5$ while we investigate $k = 1, 2, 3, 4$ of $n = 4$. There are more experimental investigations of Threshold Public Good games with *non-binary* contributions and payoff functions with two steps. For an overview see Fischbacher et al. (2011) and Normann and Rau (2015).

Our results have been summarized in (i), (ii), and (iii) above and the overview about the related literature has shown that they are novel. In particular, there is no attempt to investigate and describe behavior in the Volunteer's Dilemma, in the Stag Hunt game and in games with intermediate thresholds with a common model and the same parameters. We find that equilibrium predictions are rather successful in BTPG games. In the next section, the theory of BTPG games is presented (as far as necessary for the evaluation of the experimental results) and hypotheses connected with (i), (ii), and (iii) are formulated. Section 3 describes the experiment and Section 4 provides some descriptive statistics and a regression analysis of contribution decisions. In Section 5, a mainly equilibrium based finite mixture model is tested. Section 6 concludes.

2. Equilibria and equilibrium selection

The general theory of BTPG games is developed in Bolle (2015a). There is a set of n players $N = 1, ..., n$ who can contribute (with costs c_i) or not (without costs) to the production of a public good. If a certain threshold of contributions is surpassed, the public good is produced and the players earn G_i. The threshold is described by sufficient subsets of contributing players, here by all subsets with at least k players ($1 \leq k \leq n$). The case $0 < c_i < G_i$ is called the positive frame and $G_i < c_i < 0$ the negative frame.

Strategically neutral transformation: By renaming "contribution" as "non-contribution" (and vice versa), exchanging thresholds k and $n-k+1$, and renormalizing utilities so that "non-contribution/non-launch" has a value of zero, the negative frame is transformed into the positive frame.

In the positive frame there are $\binom{n}{k}$ pure strategy equilibria with the launch of the project where exactly k players contribute. For $k > 1$ there is one pure strategy equilibrium without the launch of the project where no one contributes. Only the latter equilibria and the "all contributing" equilibrium of the Stag Hunt game are symmetric; therefore symmetric mixed strategy equilibria of our essentially symmetric games may be assumed to be better candidates for equilibrium selection. In cases $1 < k < n$ they payoff dominate the non-contribution equilibria (Proposition 1 (iv) below).

Let us assume that the players' contribution probabilities are $p_i, i = 1, ..., n$. $p = (p_i)$ is the vector of all probabilities. $Q = Q(p)$ denotes the probability of success, i.e. that players from a sufficient subset and possibly more contribute to the production of the public good. $Q_{-i}(Q_{+i})$ denote the probability of success if i does not contribute (does contribute). These probabilities dependent only on $p_j, j \neq i$. $q_i = Q_{+i} - Q_{-i}$ is the probability that i's contribution is decisive for the production of the public good. With these definitions player i's expected revenue is

$$\begin{aligned} R_i(p) &= G_i * Q(p) - p_i c_i \\ &= G_i * [(1-p_i) * Q_{-i} + p_i * Q_{+i}] - p_i c_i \\ &= G_i * Q_{-i} + p_i * [G_i * q_i - c_i]. \end{aligned} \tag{1}$$

A mixed strategy equilibrium with $0 < p_i < 1$ requires that R_i is independent of p_i, i.e.

$$\frac{\partial R_i}{\partial p_i} = G_i * q_i - c_i = 0 \tag{2}$$

This requirement has been derived verbally by Downs (1957, p. 244) for the binary decision of voting or not. If $G_i * q_i - c_i < (>) \, 0$ then player i contributes with $p_i = 0$ (1). Inserting q_i from

(2) into (1) provides us with the equilibrium profit which i expects if he plays a mixed strategy.

$$\begin{aligned} R_i &= G_i * Q_{-i} \\ &= G_i * Q_{+i} - c_i. \end{aligned} \tag{3}$$

Proposition 1: The following statements apply in equilibrium[4]:

(i) If i plays a strictly mixed strategy, then $q_i = r_i = c_i/G_i$.
(ii) In the positive frame, $q_i > r_i$ implies $p_i = 1$ and $q_i < r_i$ implies $p_i = 0$.
(iii) In the negative frame, $q_i > r_i$ implies $p_i = 0$ and $q_i < r_i$ implies $p_i = 1$.
(iv) In equilibrium, $R_i = G_i Q_{-i}$ applies for $p_i < 1$ and $R_i = G_i Q_{+i} - c_i$ for $p_i > 0$.

Proof: (1), (2) and (3). □

If k of the n players are necessary for the production of the public good and if all $c_i/G_i = r_i = \rho$ are equal, then, in a completely mixed strategy equilibrium, all $p_i = \pi$ are equal (see Bolle, 2015a) and π is derived from

$$\rho = q_i = \binom{n-1}{k-1} \pi^{k-1} (1-\pi)^{n-k}. \tag{4}$$

For $1 < k < n$, the right hand side of (4) is a unimodal function of π with a maximum at $(k-1)/(n-1)$. Therefore (4) has either two solutions $\pi'' > \pi'$ (for small enough ρ) or one solution (border case) or no solution; i.e., completely mixed strategy equilibria do not necessarily exist and, if they exist, generically there are two. In the positive frame, the equilibrium with π'' Pareto-dominates the one with π' and vice versa in the negative frame (Proposition 1 (iv)). The remaining cases are $k = 1$ (Volunteer's Dilemma in the positive frame[5], see Diekmann, 1985) and $k = n$ (Stag Hunt game in the positive frame, see Bolle, 2015a). In these cases, unique completely mixed strategy equilibria can be explicitly determined, namely

$$p_i = 1 - \rho^{1/(n-1)} \tag{5}$$

for $k = 1$ and

$$p_i = \rho^{1/(n-1)} \tag{6}$$

for $k = n$.

Note that a completely mixed strategy equilibrium, if it exists, depends only on $\rho = c_i/G_i$ and therefore applies in the positive ($G_i > c_i > 0$) as well as in the negative ($G_i < c_i < 0$) frame[6]. Pure strategy equilibria and equilibrium selection according to Pareto-dominance, however, depend on the frame.

[4]For further implications of (2) and (3) for general BTPG games see Bolle (2015a) and Bolle (2015b).

[5]The names Volunteer's Dilemma and Stag Hunt game are chosen with a positive frame in mind. In a negative frame, $k = 4$ means that one Volunteer is needed who abstains from incurring the negative costs and thus prevents the public bad to be produced.

[6]This does not contradict the Strategically Neutral Transformation which substitutes π by $1-\pi$ and k by $n-k+1$ and thus leaves (4) unchanged.

In a pure/mixed strategy equilibrium with m pure strategies, we can apply (4) or (5) or (6) to the remaining $n' = n - m$ players, with k' instead of k indicating the remaining number of necessary contributions. For the pure strategy players Proposition 1 (ii) and (iii) must apply. Note, however, that all these strategies are asymmetric.

Let us now investigate equilibrium selection according to Harsany and Selten (HS) and risk dominance (RD). In the case of symmetric games, Harsanyi and Selten (1992) restrict their selection to the set of symmetric equilibria. These can generically be ordered according to Pareto-dominance.

Proposition 2: In a symmetric BTPG game with the threshold of k contributing players the following equilibria are selected according to the Harsanyi-Selten theory.

(i) For $k = 1$ in the positive (negative) frame (5) applies (no player contributes).
(ii) For $k = n$ in the positive (negative) frame all players contribute ((6) applies).
(iii) For $1 < k < n$ in the positive (negative) frame we get: if solutions $\pi'' \geq \pi'$ of (4) exist, then $p_i = \pi''$ (π') otherwise $p_i = 0$ (1).

Proof: (i) In the positive frame, (5) denotes the only symmetric equilibrium. In the negative frame, because of Proposition 2, the equilibrium defined by (5) yields $R_i = G_i Q_{-i} < 0$ and is therefore Pareto-dominated by the equilibrium where no one contributes. (ii) In the positive frame, Proposition 2 implies that the mixed strategy equilibrium for $k = n$ has zero payoff (because of $Q_{-i} = 0$) and is therefore Pareto-dominated by the symmetric equilibrium where all contribute with certainty ($Q_{+i} = 1$). In the negative frame, the mixed strategy equilibrium defined by (6) is the only symmetric equilibrium. (iii) For $\pi'' > \pi'$, Q_{-i} computed with π'' is larger than Q_{-i} computed with π'. Therefore, if (4) has a solution, π'' is used in the positive and π' in the negative frame. If (4) has no solution, then no one contributing (all contributing) is the only symmetric equilibrium in the positive (negative) frame. □

As an alternative selection criterion (not applied by them to symmetric games) Harsanyi and Selten (1992) define risk dominance between two equilibria p' and p'' by the "linear bicentric tracing procedure" where there is a prior probability $1-t$ that player i's co-players $j \neq i$ choose "contributing" with a probability which is a 0.5-mixture of p'_j and p''_j, and a prior probability t that they play an equilibrium strategy (under consideration of the priors). For $t = 0$ every player plays a best reply to others' 0.5-mixtures and for $t = 1$ all players play an equilibrium of the game. If there is a path of equilibria with $0 \leq t \leq 1$ connecting the generically unique equilibrium with $t = 0$ with p' (for $t = 1$) then p' is said to risk dominate p''. Because it is necessary to distinguish several cases it is tedious to characterize the risk dominance relations between the symmetric equilibria for all symmetric BTPG games. Therefore we concentrate here on the positive frame of our experimental games which are characterized by $n = 4$ and $\rho = 0.4$. S0 denotes the pure strategy equilibrium (for $k = 2, 3, 4$) where no one contributes, S1 denotes the pure strategy equilibrium (for $k = 4$) where all contribute, and p^k denotes the equilibrium defined by π'' from (4). Note that, in the case $k = 1$, we have a unique symmetric equilibrium.

Proposition 3: Let us assume the positive frame of symmetric BTPG games with $n = 4$ and $\rho = 0.4$. Then S0 risk dominates S1 ($k = 4$) and p^3 and p^2.

Proof: Because of $q_i(0.5) = \binom{n-1}{k-1} 0.5^{n-1} < \rho$ in all cases, the best reply of player i to other players playing $p_j = 0.5$ ($k = 4$) or $p_j < 0.5$ ($k = 2$ and 3) is $p_i = 0$. This coincides with S0 and is therefore selected by the linear tracing procedure (Lemma 4.17.7 in Harsanyi and Selten, 1992). □

Propositions 2 and 3 provide, for the positive frame with $k > 1$, contradicting predictions from the selection principles payoff dominance (HS) and risk dominance (RD). The same applies for the negative frame in the cases $k < 4$ where risk dominance predicts that all players contribute.

The introduction of *altruistic* and/or *warm glow* players in the spirit of Andreoni (1989, 1990) suggestion changes the game only insofar as the cost/benefit relations r_i are multiplied by a factor. Following Andreoni (1990) we add an "altruistic" term by substituting G_i by $G_i + a_i * G_{-i}$ with $G_{-i} = \sum_{j \neq i} G_j$ and we introduce an additional "warm glow" utility $b_i * c_i$ of contributing to the public good. With such a utility function, players who play mixed strategies with probabilities p_i have revenues

$$R_i = Q * (G_i + a_i * G_{-i}) - (1 - b_i) * p_i * c_i \tag{7}$$

This results in the equivalent to (2),

$$(G_i + a_i * G_{-i}) * q_i - (1 - b_i) * c_i = 0 \tag{8}$$

Proposition 5: The introduction of altruistic and/or warm glow players results in an equilibrium condition for mixed strategies

$$q_i = r_i * s_i \qquad \text{with} \quad s_i = \frac{G_i(1 - b_i)}{G_i + a_i * G_{-i}} \tag{9}$$

Proof: (8). □

For $a_i = 0$ we get $s_i = 1 - b_i$ which is independent of the benefits G_i. Below, we will assume $a_i = a = 0$, $b_i = b$ and estimate a unique $s_i = s = 1 - b$. We do not use the option of estimating two different s for average $a \neq 0$ and b according to our two different benefit values because the model estimation is already "perfect".

The following hypotheses will be tested. Every player plays games with eight repetitions in a stranger design and is characterized by his individual contribution frequency ICF, a number between 0 and 8. Aggregate behavior of player types in a game described by a threshold k, costs c_i, and benefits G_i can be characterized by the frequency distribution $f(\text{ICF}; k, c_i, G_i)$ or, more aggregated, by the average contribution probability $\text{ACP}(k, c_i, G_i)$ of this group of players.

(H1) Comparisons within the *same frame*: For $\lambda > 0$, (a) $\text{ACP}(k, c_i, G_i) = \text{ACP}(k, \lambda c_i, \lambda G_i)$ and (b) $f(\text{ICF}; k, \lambda c_i, \lambda G_i) = f(\text{ICF}; k, c_i, G_i)$.

(H2) Comparison *between frames*: (a) $\text{ACP}(k, c_i, G_i) = 1 - \text{ACP}(n - k + 1, -c_i, -G_i)$ and (b) $f(\text{ICF}; k, c_i, G_i) = 8 - f(\text{ICF}; n - k + 1, -c_i, -G_i)$

(H3) Behavior can be described by the equilibrium selected by HS.

(H1) is motivated by Proposition 1 and (H2) by the Strategically Neutral Transformation (see above). HS is described in Proposition 2. As (H3) is a rather strong hypothesis we will test also implications and a weaker version.

(H4) ACPs are closer to the HS prediction than to the risk dominance (RD) prediction.

(H5) With $c_i/G_i = 0.4$, $\text{ACP}(k, c_i, G_i)$ increases with k.

(H3') Behavior can be described as a mixture of mainly equilibrium behavior.

RD is described in Proposition 3. (H5) stems from applying (4), (5) and (6); see Table II. The elements of the mixture in (H3') will be described below.

3. Experiments

All our experimental games are with four players and $r_i = c_i/G_i = 0.4$. In Treatment 1 (positive frame), players 1 and 3 with $(c_i, G_i) = (4, 10)$ are called small players; players 2 and 4 with $(c_i, G_i) = (8, 20)$ are called large players. In Treatment 2 (negative frame), G_i and c_i have the same absolute values but are both negative.

We have conducted a computerized laboratory experiment (implemented in a z-tree program design, Fischbacher, 2007) with 8 subjects per session who were randomly assigned to a player type which they kept during the whole session. In every session, there were four small and four large players. In every period, subjects were randomly allocated to two groups with the restriction that the groups consisted of two small and two large players. During eight periods the threshold k was the same; then it was changed and again kept during the next eight periods, and so on. During 32 periods all thresholds were adopted in a random order but with the restriction that, across the 10 sessions, each k was played either 2 or 3 times at each of the four positions and each k was played five times in the first two positions (periods 1-16) and five times in the last two positions (periods 17-32). Subjects were not informed about the order of the thresholds in the beginning, but only when the threshold was changed. We used a stranger design, i.e. the composition of the groups was changed after each round and the co-players could not be identified. Subjects were informed about how many players contributed to the public good but not who contributed. Hence, players were unable to build a reputation. 10 sessions in both treatments and eight subjects per session resulted in a total of 160 subjects. Before subjects played the BTPG games, they were given printed instructions and had the possibility to ask

Table I: Game parameters

	Player type	Endowment E_i	Contribution costs c_i	Benefits G_i	Expected revenue ($p_i = 0.5$)
Treatment 1	small	8	4	10	11
(positive frame)	large	8	8	20	17
Treatment 2	small	20	-4	-10	17
(negative frame)	large	20	-8	-20	14

Values in lab dollars in the two treatments and average expected revenues if all players always play mixed strategies with $p_i = 0.5$.

questions. Instructions contained general information, the description of the threshold public good game and two example calculations. Furthermore, they had to answer five comprehension questions to make sure that everyone understood the game. The experiment did not start before all subjects had answered the questions correctly. In cases of problems, personal advice was given. In every period, the subjects were reminded of the actual threshold and, every 8th period, the changing of the threshold was announced. In each period subjects were informed on the *decision screen* that the group composition had been changed and they were required to decide whether or not to contribute. On the *profit display screen* they were informed about the number of contributing players and whether the threshold was reached. They further received information about their payoff in the current period.

Each round, players were endowed with an amount of lab dollars E_i. Table I shows the parameters of all player types in all treatments. For each lab dollar earned, subjects were paid 4 Eurocents. After the experiment, subjects were presented three incentivized questions testing their understanding of probability calculus. For each correctly answered question (on average two), the subject was paid one additional Euro. Participants earned between 14 and 33 Euros with an average of 23.29 Euros. Sessions lasted roughly 45 minutes.

4. Results

4.1. Average contribution probabilities

In Table II, average contribution frequencies ACPs are reported. Tests are carried out with respect to our hypotheses. We compare small and large players, the positive and the negative frame, the influence of k, and the relative performance of HS vs. RA.

Result 1: *Small and large players show similar ACPs except for* $k = 4$. (H1) (a) is rejected for $k = 4$ in both treatments; it is not rejected for other thresholds.

Result 2: *ACPs in the positive and the negative frame are mirrored,* i.e. $\text{ACP}(k, c_i, G_i) = 1 - \text{ACP}(5 - k, -c_i, -G_i)$. (H2) (a) is not rejected in any of the eight comparisons.

We find that ACPs increase from k to $k + 1$, i.e., the alternative to (H5), non-increasing ACPs, is significantly rejected in 9 of 12 one-sided Wilcoxon matched pairs tests. In the positive frame,

Table II: Average contribution probabilities (ACPs)

Threshold	Positive frame		Negative frame		HS	
k	small player	large player	small player	large player	positive	negative
1	0.35*	0.37*	0.30	0.26	0.26	0
2	0.49*	0.56	0.43*	0.39*	0.46	0.22
3	0.61*	0.63*	0.57*	0.49*	0.76	0.54
4	0.74**	0.81	0.75**	0.59	1	0.74

Average contribution probabilities (ACPs) in Treatments 1 (positive frame) and 2 (negative frame). Small player type S with $(G_S, c_S) = (10, 4)$ and large player type L with $(G_L, c_L) = (10, 4)$ in the positive frame. HS: equilibrium according to Harsanyi and Selten (1992) as described in Proposition 3 and without altruism or warm glow. ** indicates two-sided Wilcoxon matched-pairs test for small vs. large players. * indicates one-sided Wilcoxon matched pairs-test of non-increasing ACPs for k vs. $k+1$. No significant results in two-sided Mann-Whitney tests between ACP(k, c_i, G_i) and $1-$ACP$(5-k, -c_i, -G_i)$. All tests are based averages from 10 sessions and $p < 0.05$.

this means an attempt to meet the increasing threshold, which is apparently stronger than the fear to waste one's contribution if others do not cooperate. In the negative frame, these arguments are mirrored.

Result 3: *In both treatments, ACPs increase with the threshold k.* For the transition from k to $k+1$, the alternative to (H5), namely decreasing or constant ACPs, is rejected in 9 of 12 tests.

At last, we compare the predictions of payoff dominance with those of risk dominance (RD) in the cases $k > 1$ in the positive frame and $k < 4$ in the negative game. The HS (payoff dominance) predictions are reported in Table II. RD predicts zero contributions in the positive frame with $k > 1$ and full contributions in the negative frame with $k < 4$. We have conducted one-sided Wilcoxon matched pairs tests to the alternative to (H4), smaller or equal differences between ACPs and RD than between ACPs and HS. The alternative is rejected in 15 of 16 cases (small and large players, positive and negative frame), 13 times on the 1% level and two times on the 5% level. The tests are based again on the average ACPs in the ten sessions.

Result 4: *In both treatments, ACPs are closer to HS than to RD.* The alternative to (H4), smaller or equal differences between ACPs and RD than between ACPs and HS, is rejected in 15 of 16 cases on the 5% level.

This result is particularly interesting for the Stag Hunt game ($k = 4$ in the positive, $k = 1$ in the negative frame) because, in the literature, the comparison is predominantly made in this game and most experimental studies favor RD. We find that the performance of HS even improves in the course of time. ACPs are rather closer to full (zero in the negative frame) contributions if we concentrate on periods 17–32 (see Table VII in the appendix) where we find 90% contributions for $k = 4$ in the positive frame (average of small and large players) and 6% for $k = 1$ in the negative frame.

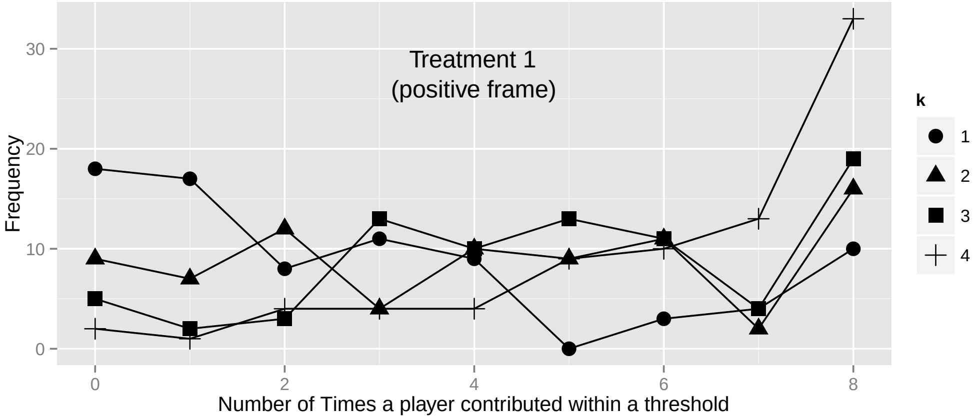

Figure 1: Frequency distribution of individual contribution frequencies (ICFs) in Treatment 1 (positive frame). k = threshold. For every k, 8 decisions by 80 individuals.

4.2. The distribution of individual contribution frequencies

Every decision situation occurs 8 times so that every subject can contribute to the public good (for every threshold) between 0 and 8 times. We call this number the *individual contribution frequency* (ICF). The distributions of these ICFs are provided in the appendix for every player type and for the cases that a game had been one of the first two games (periods 1-16) or one of the last two games (periods 17-32). They are not only used for second tests of (H1) and (H2) but provide also the data for the estimation of the finite mixture model in the next section. In Figures 1 and 2 aggregations over player types are presented. The question of whether there is a trend in the decisions is investigated in a regression analysis below. The individuals with 0 or 8 contributions can be assumed to play a pure strategy or to use mixture probabilities close to 0 or 1. According to this criterion, 40% of our subjects play (almost) pure strategies in single games, 14.3% with zero contributions and 25.8% with full contributions. The fact that the extreme ICFs in the cases $k = 1$ and $k = 4$ are accompanied by high numbers in the neighboring frequency classes 1 or 7 is most plausible if many subjects use an "almost pure" strategy with contribution probabilities close to 0 or 1.

We test (H1) (b) by comparing the ICFs of small and large players (from all periods) in a chi-square test, i.e., we compare the frequencies of ICF $= 0, 1, ..., 8$ for $k = 1, 2, 3, 4$. In the positive frame we get $\chi^2 = 31.8$(df $= 31$[7], $p = 0.42$), in the negative frame $\chi^2 = 34.3$ (df $= 32$, $p = 0.36$).

Result 5: *Players in the same frame with the same cost/benefit ratio show the same distribution of ICFs.* (H1) (b) is not rejected.

At first glance we notice that the frequency of ICF $= 8$ in Treatment 1 with the threshold $k = 4$ meets that of ICF $= 0$ in Treatment 2 with the threshold $k = 1$. If behaviors in Treatments 1 and 2 are completely mirrored, we should find these similarities also for other ICFs. A chi-square

[7]There is one class ($k = 1$, ICF $= 6$) with zero contributions which is united with a neighboring class.

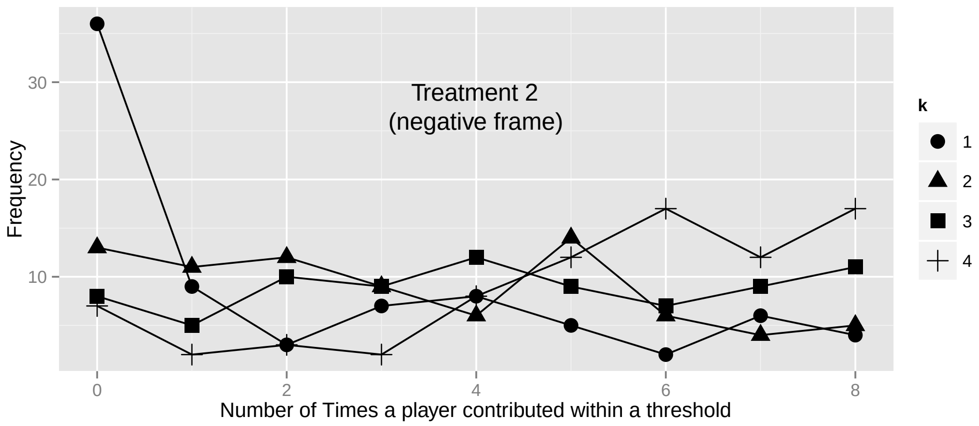

Figure 2: Frequency distribution of individual contribution frequencies (ICFs) in Treatment 2 (negative frame). k = threshold. For every k, 8 decisions by 80 individuals.

test ($\chi^2 = 34.4, \mathrm{df} = 32, p = 0.35$) of this hypothesis, based on the frequencies in Figures 1 and 2, does not indicate significant differences.

Result 6: *Behavior in negative frame is equal to the "mirrored" behavior in the positive frame.* (H2) (b) is not rejected.

4.3. Regression analysis

Table III shows the results of a regression analysis which takes into account the clusters of subjects who participated in the same session. We analyze the data from the first ($t < 17$) and the second half ($t > 16$) of the experiments separately because learning effects might exist. At least, the coefficients of k are considerably larger in $t > 16$. In Section 5, the difference between the first and the last periods turns out to be crucial for the rejection (or not) of the finite mixture model. Stylized results of our regression analysis are, first, confirmations of previous results, namely:

- The dummy for the large player is insignificant, i.e. (i) and Results 1 and 5 are supported.
- The "mirror image" character of Treatments 1 and 2 is confirmed because absolute coefficients of PivotalPlayerLastRound, and Threshold k are quite similar, i.e. (ii) and Results 2 and 6 are supported.
- The higher the threshold, the more frequent are contributions, i.e. Result 3 is supported.

There is, however, one new insight from this analysis. PivotalPlayerLastPeriod is a dummy variable that indicates whether a player's previous decision was crucial for reaching / not reaching the threshold.

Result 7: *Contributing is more probable if a player has been critical in the last round.*

Table III: Logit-Regressions

Variable	Periods 2 – 32 Treatment 1 (pos. frame)	Periods 2 – 32 Treatment 2 (neg. frame)	Periods 2 – 16 Treatment 1 (pos. frame)	Periods 2 – 16 Treatment 2 (neg. frame)	Periods 17 – 32 Treatment 1 (pos. frame)	Periods 17 – 32 Treatment 2 (neg. frame)
Intercept	-1.69^*	-0.55^*	-1.11^*	-0.16	-1.45^*	-2.33^*
	(0.28)	(0.26)	(0.36)	(0.28)	(0.64)	(0.57)
PivotalPlayer	0.88^*	-1.00^*	0.69^*	-0.96^*	0.96^*	0.79^*
LastPeriod	(0.12)	(0.10)	(0.15)	(0.13)	(0.15)	(0.17)
LargePlayer	0.23	-0.28	0.12	-0.18	0.37	-0.47
	(0.21)	(0.20)	(0.22)	(0.22)	(0.30)	(0.27)
Threshold k	0.55^*	0.51^*	0.38^*	0.21^*	0.76^*	0.92^*
	(0.07)	(0.07)	(0.10)	(0.08)	(0.11)	(0.12)
Period	0.01	-0.02^*	0.00	0.02	-0.02	0.01
	(0.007)	(0.006)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	2480	2480	1200	1200	1280	1280
$-\log L$	1515	1526	788	789	710	693

Logit-Regressions of contribution decisions with standard errors in parentheses. * indicates significance at 5% level.

Moreover, we conducted several alternative regressions to control for various influences. We used sex and study and answers to incentivized questions about probability calculus as controls and we used the rounds (from 1 to 8) within a threshold level. We alternatively also conducted all regressions with an autoregressive term. None of these variations changes the outcome of the analysis.

5. A finite mixture model of behavior

From Figures 1 and 2 we know that players in a certain game (described by treatment and threshold) use different strategies. Some play pure or, more probably, almost pure strategies; others possibly play mixed strategies with constant mixture probabilities (or switch between different strategies). In this situation we cannot successfully describe behavior without introducing different modes of play.

In addition to the selection principle **HS**, we introduce four possible deviations concerning other modes of play and, in addition, small random deviations. Ideal for the support of HS would be that the share of HS is estimated to be large and the random deviations to be small. The four other selections are:

- **FE** (false equilibrium) selects the equilibrium selected by HS for the "mirrored game".
- **E** selects an efficient (group income maximizing) equilibrium.
- **S0** where $p_i = 0$ is played in the positive frame and $p_i = 1$ in the negative frame.
- **S1** where $p_i = 1$ is played in the positive frame and $p_i = 0$ in the negative frame.

We also allow for average warm glow by multiplying all cost/benefit ratios $r_i = 0.4$ with a

Table IV: Selected Equilibria

k	E	HS	FE
1	$p = (p_E, p_E, 0, 0)$	(5) with $\rho = 0.4s$	$p = (0, 0, 0, 0)$
2	$p = (1, 1, 0, 0,)$	π'' from (4) with $\rho = 0.4s$	π' from (4) with $\rho = 0.4s$
3	$p = (1, 1, p_E, p_E)$	π'' from (4) with $\rho = 0.4s$	π' from (4) with $\rho = 0.4s$
4	$p = (1, 1, 1, 1)$	$p = (1, 1, 1, 1)$	(6) with $\rho = 0.4s$

Selected equilibria in Treatment 1 according to the two main principles E and HS and the "error" principle FE. (Treatment 2 correspondingly)

constant factor s as described in (9). This influences the equilibria selected according to HS and FE.

In the positive (negative) frame, S0 is an equilibrium for $k = 2, 3, 4$ (for $k = 1, 2, 3$) and S1 is an equilibrium for $k = 4$ (for $k = 1$). S1 may be chosen by some subjects because they try to follow a social norm or because of their extremely high warm glow (low s_i). S0 coincides with the selection of a symmetric equilibrium according to risk dominance **RD** except in the games with $k = 1$ in the positive frame and $k = 4$ in the negative frame where there is only one symmetric equilibrium which is then selected. We can substitute S0 by RD without loss of fit (see Table VI); therefore it is a question of taste whether S0 or RD should be used for the finite mixture model. RD predicts only equilibrium behavior but, on the other hand, S0 is the natural counterpart of S1. We finally give more weight to the latter argument and proceed with S0.

E predicts asymmetric equilibria except for $k = 4$ in the positive ($k = 1$ in the negative) frame. The problem with E in the positive frame (and accordingly in the negative frame) is that, under the threshold $k = 1$, it is not determined who of the two small players should contribute and, under $k = 3$ where the two small players contribute with certainty, who of the large players should contribute. We fill this gap by assuming that, in both situations, the two players play a mixed strategy with the probability p_E while players 2 and 4 do not contribute ($k = 1$) or players 1 and 3 always contribute ($k = 3$). This probability may be equal to the mixed strategy probability of the Volunteer's Dilemma with two players, but in order to give E a better chance we estimate p_E from our data. Of course E is no longer efficient in these cases but Proposition 1 (iv) shows that a thus defined E is "more efficient" than the symmetric completely mixed strategy equilibrium in HS. If the two mixed strategy players play the Volunteer's Dilemma mixture in E, they earn $G_i - c_i$ because $Q_{+i} = 1$. Players who do not contribute earn more. Therefore, in the case $k = 1$, players 1 and 3 in E earn the same as in HS and players 2 and 4 earn more. In the case $k = 3$, players 2 and 4 earn more in E because $Q_{+i} < 1$ in HS. Computations for players 1 and 3 result in small losses $Q_{+i} = 0.84$ in E compared with $Q_{+i} = 0.88$ in HS; but the expected sum of revenues (social product) is larger in E.

Instead of establishing a Random Utility model we simply add a constant probability ϵ for a random deviation from the binary decisions of the players. The main effect of ϵ is that pure strategy equilibria are now played with $p_i = \epsilon$ instead of $p_i = 0$ and $p_i = 1-\epsilon$ instead of $p_i = 1$. In the mixed strategy equilibria, instead of the computed equilibrium probabilities p_i, we apply $p_i * (1-\epsilon) + (1-p_i) * \epsilon$. Every selected mode of play leads to a binomial distribution of ICFs of a player in the eight rounds of a game. Shares α_E, $\alpha_S P$, $\alpha_F E$, α_1, and α_0 determine an according

Table V: Finite mixture parameter estimations

per	α_{HS}	α_{FE}	α_E	α_1	α_0	S	ϵ	p_E	χ^2	df	p
all	0.405	0.334	0.001	0.165	0.095	0.687	0.069	0.578	171.1	121	0.003
	(0.018)	(0.016)	(0.010)	(0.014)	–	(0.025)	(0.005)	(0.552)			
<17	0.303	0.363	0.001	0.168	0.162	0.556	0.134	0.521	173.9	121	0.002
	(0.041)	(0.033)	(0.075)	(0.031)	–	(0.046)	(0.009)	(0.555)			
>16	0.490	0.279	0.011	0.170	0.062	0.735	0.035	0.779	142.7	121	0.087
par7	(0.025)	(0.023)	(0.022)	(0.020)	–	(0.028)	(0.005)	(0.242)			
>16	0.492	0.283	–	0.173	0.042	0.743	0.035	–	142.8	123	0.107
par5	(0.024)	(0.021)	–	(0.019)	–	(0.122)	(0.005)	–			

Parameter estimations for Treatment 1 and 2 and all thresholds, α_{HS}, α_{FE}, α_E, α_1 and $\alpha_0 = 1-\alpha_{HS}-\alpha_{FE}-\alpha_E-\alpha_1$ are the shares of HS, FE, E, S1 and S0. Last row: $\alpha_E = 0$ (then p_E is irrelevant).

mixture of binomial distributions. These will be compared with the empirical ICFs of players (see Table VII in the appendix).

We test the model jointly (with the same parameters) for Treatments 1 and 2 and for the four thresholds. We estimate the seven parameters $\alpha_E, \alpha_{SP}, \alpha_{FE}, \alpha_1, (\alpha_0 = 1-\alpha_{SP}-\alpha_{FE}-\alpha_1), \epsilon, S$ and p_E. The α-parameters denote the shares of the respective behavioral modes. For the parameter estimation, we apply the χ^2 minimum method[8] with data from the 144 frequencies in Table VII in the appendix. In order to control for learning, we also estimate parameters based on the decisions in the first half (periods 1-16) or the second half (periods 17-32). Every threshold k was played in five of the ten sessions as one of the first two games, i.e. in periods 1-16. The number of degrees of freedom is df $= 144-4*4-7 = 121$. In Table V we present the estimations.

Only the late decisions are successfully predicted by the model, i.e. without rejection in the χ^2 test on the 5% level. Half of our subjects play the strategy selected by HS, a quarter play FE, one in six contributes almost always in the positive and almost never in the negative frame, and one in sixteen contributes almost never in the positive and almost always in positive frame. As expected, ϵ is rather small (0.035). Surprisingly, according to this model, nobody plays E. $p = 0.087$ indicates that, at least on the basis of such a χ^2 test, no significantly better description of the results is possible. It is, however, possible to remove E from this estimation. Without the two parameters α_E and p_E, χ^2 increases by only 0.1 points and, because of the increased number of degrees of freedom, the p value even increased slightly. Further disregarding modes of behavior or setting $s = 1$ (no altruism or warm glow) or $\epsilon = 0$ leads to rejection of the simplified model (see Table VI). We can substitute S0 by RD which changes neither the fit of the model nor the parameters. While the share of S0 is estimated as 6.6%, RD's share is estimated as 6.3%.

[8]See Newey and West (1987) and Berkson (1980). The method is asymptotically equivalent to maximum likelihood. If we estimate the par5 model (Table V) by maximum likelihood, then $-\log L$ decreases by only two points to 585.8 and the estimated parameters are minimally affected. Chi-square allows the absolute evaluation of a model while $-\log L$ facilitates the comparison of models. For the numerical computations the methods "L-BFGS-B" and "Nelder-Mead" from the R-library have been used. Standard errors are computed by the inverse of the Hessian.

Table VI: Finite mixture parameters after simplification

Model	par5	$s = 1$	$\alpha_{FE} = 0$	$\alpha_0 = 0$	$\alpha_1 = 0$	$\epsilon = 0$	RD
χ^2	142.8	189.0	230.2	175.4	442.7	252.0	142.9
df	123	124	124	124	124	124	123
$P(\chi^2)$	0.107	0.0002	$2 * 10^{-8}$	0.002	0	10^{-10}	0.106
$-\log(L)$	588.6	601.4	622.7	599.9	701.4	604.9	588.1

Fit after further simplifications of the five-parameter model par5 (periods>16) of Table VII and after substituting S0 by risk dominance (RD). The log-likelihood scores are not comparable with those of the regression where single decisions are evaluated. Here the logarithms of mixtures of binomial distributions determine $-\log L$ which is systematically comparable with the χ^2 score. If we determine $-\log L$ by the probability of single decisions, we get 1469 (BIC=2966) for the estimation of the five-parameter model while the score in the regression is $710 + 693 = 1403$ (BIC=2892). Therefore two regressions for Treatments 1 and 2 show a better joint fit than the common finite mixture model for the two treatments.

When comparing the estimations for the first 16 and the last 16 periods (Table V), we may conclude that, in the beginning, the subjects were strongly influenced by their costs of contributing. In the positive frame they wanted to avoid costs and in the negative frame they wanted to collect the negative costs. This leads to a mode of play similar to FE or S0, i.e. playing the Pareto-inferior less-contribution or even no contribution equilibria in Treatment 1 (Pareto-inferior more-contribution or full contribution equilibria in Treatment 2). During the first 16 periods, learning towards HS occurred so that we have no stable modes of behavior during the first 16 periods with the consequence that the model is rejected. During the second half of the experiment behavior was apparently stable enough so that the model cannot be rejected on the basis of a χ^2 test. The remaining FE-players are still overly influenced by their costs or they wanted to play HS but confused it with FE.

Note that the regression model (Table III) and the finite mixture model (Table V) do not really compete but are complementary. The regression model tests for influences as the pivotality of a player's previous decision which are absent in the finite mixture model, and the finite mixture model allows highly non-linear systematic influences of the threshold while the regression model describes only linear ones.

6. Conclusion

BTPG is an important class of games with many applications, such as forming teams for producing a public good/ preventing a public bad or also costly voting in committees and parliament. The most severe obstacle for the application of theory is the plethora of equilibria and the question of their relative and absolute importance. Concerning equilibrium play we have three main results. (i) As theory predicts, if the cost/benefit ratio is the same, then subjects play (almost) the same mixture of strategies. (ii) After switching from a positive to a negative frame, the theoretically expected "mirrored" behavior can be observed. This is not obvious because there are many examples where another frame (for example a common pool frame instead of a public good frame) induces significantly different behavior.

Our main result (iii) concerns the question whether behavior can be described by a common model with mainly equilibrium play. Risk dominance arguments are rejected in favor of payoff dominance. For efficient behavior, a zero share in a finite mixture model is estimated. HS performs best, especially after some learning when its share is estimated to be 49% in the finite mixture model which cannot be rejected in a chi-square test. So, HS has experienced moderate support.

All these results mean progress for our understanding of how BTPG games are played, because (i) and (iii) are novel and (ii) shows that framing effects need not be ubiquitous. (iii) is particularly important because it is shown that behavior in the Volunteer's Dilemma, the Stag Hunt game, and in games with intermediate thresholds can be explained with a common model with the same parameters. Experimental investigations have coped so far only with one or two thresholds. Of course, many questions remain, in particular how people behave in mildly and highly asymmetric situations as well as, for example, in situations with incomplete information. In the face of the rather limited number of experimental investigations it is too early for conclusions concerning behavior outside the laboratory. The message, however, of the supplementary description of behavior by a regression and a finite mixture model with (mostly) equilibrium play is that equilibria as well as beliefs (different from rational expectations as in game theoretic equilibria) about the average play of others seem both to be important for a good description of behavior.

Acknowledgment

We are grateful for the funding by the German Ministry of Education and Research (FKZ 01LA1139A) and the Deutsche Forschungsgemeinschaft (project BO 747/14-1)

References

Al-Ubaydli, O., Jones, G., and Weel, J. (2013). Patience, cognitive skill, and coordination in the repeated stag hunt. *Journal of Neuroscience, Psychology, and Economics*, 6(2):71.

Andreoni, J. (1989). Giving with impure altruism: applications to charity and ricardian equivalence. *The Journal of Political Economy*, pages 1447–1458.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm glow giving. *Economic Journal*, pages 464–477.

Andreoni, J. (1995). Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics*, pages 1–21.

Bartling, B., Fischbacher, U., and Schudy, S. (2015). Pivotality and responsibility attribution in sequential voting. *Journal of Public Economics*.

Battalio, R., Samuelson, L., and Van Huyck, J. (2001). Optimization incentives and coordination failure in laboratory stag hunt games. *Econometrica*, pages 749–764.

Berkson, J. (1980). Minimum chi-square, not maximum likelihood! *The Annals of Statistics*, pages 457–487.

Bolle, F. (2015a). Costly voting - a general theory of binary threshold public goods. Discussion Paper, Frankfurt (Oder).

Bolle, F. (2015b). A note on payoff equivalence of the volunteer's dilemma and the stag hunt game and inferiority of intermediate thresholds. *International Game Theory Review*, page 1550004.

Breitmoser, Y., Tan, J., and Bolle, F. (2015). Voluntary contributions by consent or dissent. *Games and Economic Behavior*, page forthcoming.

Bullinger, A. F., Wyman, E., Melis, A. P., and Tomasello, M. (2011). Coordination of chimpanzees (pan troglodytes) in a stag hunt game. *International Journal of Primatology*, 32(6):1296–1310.

Büyükboyacı, M. (2014). Risk attitudes and the stag-hunt game. *Economics Letters*, 124(3):323–325.

Chen, X.-P., Au, W. T., and Komorita, S. (1996). Sequential choice in a step-level public goods dilemma: The effects of criticality and uncertainty. *Organizational Behavior and Human Decision Processes*, 65(1):37–47.

Dawes, R. M., Orbell, J. M., Simmons, R. T., and Van De Kragt, A. J. (1986). Organizing groups for collective action. *American Political Science Review*, 80(04):1171–1185.

Diekmann, A. (1985). Volunteer's dilemma. *Journal of Conflict Resolution*, pages 605–610.

Diekmann, A. (1993). Cooperation in an asymmetric volunteer's dilemma game theory and experimental evidence. *International Journal of Game Theory*, 22(1):75–85.

Diekmann, A. and Przepiorka, W. (2015). Punitive preferences, monetary incentives and tacit coordination in the punishment of defectors promote cooperation in humans. *Scientific reports*, 5.

Downs, A. (1957). An economic theory of political action in a democracy. *The journal of political economy*, pages 135–150.

Dufwenberg, M., Gächter, S., and Hennig-Schmidt, H. (2011). The framing of games and the psychology of play. *Games and Economic Behavior*, 73(2):459–478.

Engelmann, D. and Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American economic review*, pages 857–869.

Erev, I. and Rapoport, A. (1990). Provision of step-level public goods the sequential contribution mechanism. *Journal of Conflict Resolution*, 34(3):401–425.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868):137–140.

Feltovich, N., Grossman, P. J., et al. (2013). The effect of group size and cheap talk in the multi-player stag hunt: experimental evidence. Technical report, Monash University, Department of Economics.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*, 10(2):171 – 178.

Fischbacher, U., Güth, W., and Levati, M. V. (2011). Crossing the point of no return: A public goods experiment. Technical report, Jena economic research papers.

Franzen, A. (1995). Group size and one-shot collective action. *Rationality and Society*, 7(2):183–200.

Goeree, J. K., Holt, C. A., and Moore, A. K. (2005). An experimental examination of the volunteer's dilemma. *Unpublished working paper*.

Goren, H., Kurzban, R., and Rapoport, A. (2003). Social loafing vs. social enhancement: Public goods provisioning in real-time with irrevocable commitments. *Organizational Behavior and Human Decision Processes*, 90(2):277–290.

Harsanyi, J. C. and Selten, R. (1992). *A General Theory of Equilibrium Selection in Games*, volume 1 of *MIT Press Books, 2nd edition*. The MIT Press.

Kritikos, A. and Bolle, F. (2001). Distributional concerns: equity-or efficiency-oriented? *Economics Letters*, 73(3):333–338.

McEvoy, D. M. (2010). Not it: opting out of voluntary coalitions that provide a public good. *Public Choice*, 142(1-2):9–23.

Newey, W. K. and West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, pages 777–787.

Normann, H.-T. and Rau, H. A. (2015). Simultaneous and sequential contributions to step-level public goods one versus two provision levels. *Journal of Conflict Resolution*, 59(7):1273–1300.

Palfrey, T. R. and Rosenthal, H. (1991). Testing for effects of cheap talk in a public goods game with private information. *Games and economic behavior*, 3(2):183–220.

Park, E.-S. (2000). Warm-glow versus cold-prickle: a further experimental study of framing effects on free-riding. *Journal of Economic Behavior & Organization*, 43(4):405–421.

Przepiorka, W. and Diekmann, A. (2013). Individual heterogeneity and costly punishment: a volunteer's dilemma. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1759):20130247.

Rose, S. K., Clark, J., Poe, G. L., Rondeau, D., and Schulze, W. D. (2002). The private provision of public goods: tests of a provision point mechanism for funding green power programs. *Resource and energy economics*, 24(1):131–155.

Rousseau, J.-J. (1997). *Rousseau:'The Social Contract'and Other Later Political Writings*. Cambridge University Press. trans. Victor Gourevitch.

Rydval, O. and Ortmann, A. (2005). Loss avoidance as selection principle: evidence from simple stag-hunt games. *Economics Letters*, 88(1):101–107.

Sonnemans, J., Schram, A., and Offerman, T. (1998). Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior & Organization*, 34(1):143 – 161.

Van de Kragt, A. J., Orbell, J. M., and Dawes, R. M. (1983). The minimal contributing set as a solution to public goods problems. *American Political Science Review*, 77(01):112–122.

Van Huyck, J. B., Battalio, R. C., and Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review*, pages 234–248.

Whiteman, M. A. and Scholz, J. T. (2010). Social capital in coordination experiments: Risk, trust and position.

Willinger, M. and Ziegelmeyer, A. (1999). Framing and cooperation in public good games: an experiment with an interior solution. *Economics letters*, 65(3):323–328.

A. Data

Table VII: Frequency of individual contribution frequencies (ICF)

		Periods 1 – 16									Periods 17 – 32								
k		0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
	Treatment 1																		
	small players																		
1		4	5	1	2	4	0	1	1	2	3	7	2	2	3	0	0	0	3
2		1	1	4	2	2	1	2	0	3	4	3	4	0	3	3	1	0	6
3		1	0	1	5	3	5	3	1	5	1	1	2	3	0	1	3	1	4
4		2	0	2	1	3	3	3	4	2	0	0	0	0	0	3	4	1	12
	large players																		
1		6	1	3	4	1	0	2	2	1	5	4	2	3	1	0	0	1	4
2		1	2	2	1	2	3	1	1	3	3	1	2	1	3	2	7	1	4
3		1	1	0	4	5	4	3	1	5	2	0	0	1	2	3	2	1	5
4		0	1	2	3	1	3	2	4	4	0	0	0	0	0	0	1	4	15
	Treatment 2																		
	small players																		
1		2	2	0	1	7	2	1	1	4	16	2	1	1	0	0	0	0	0
2		4	4	0	1	2	4	2	1	2	2	1	6	3	2	3	1	1	1
3		3	0	2	2	3	1	3	2	4	1	2	1	3	3	4	0	3	3
4		2	0	0	0	4	2	6	3	3	1	0	0	1	0	0	4	6	8
	large players																		
1		4	0	1	5	1	3	1	5	0	14	5	1	0	0	0	0	0	0
2		2	2	5	3	2	4	0	1	1	5	4	1	2	0	3	3	1	1
3		1	1	5	2	3	2	3	3	0	3	2	2	2	3	2	1	1	4
4		0	1	1	1	4	4	5	2	2	4	1	2	0	0	6	2	1	4

Individual contribution frequencies (ICF) by threshold, player type, treatment and whether the threshold was played in the first half (period < 17) or in the second half (period > 16).

B. Example Instructions[9]

Welcome

You are participating in an economic experiment. You will receive your payoff personally and directly after the experiment. The payoff depends on your own decisions and the decisions of your co-players.

Please, turn off your cellphone and similar devices. The entire experiment is conducted on the computer. During the course of the experiment, please do not speak and do not communicate with other participants in any other way.

Below you will find an explanation of the experiment. Please read it carefully. If you have questions notify the experimenter. The experimenter will then answer them. After reading these instructions you will answer several test questions. If you have problems answering these questions, please also notify the experimenter.

Instructions

- In this experiment you have to make decisions in several periods.
- In each period groups of 4 players are built. You are always player 1 in your group.
- Each period each player is endowed with 8 points.
- Each player can either choose A or B.
- For now choosing B has no impact on your points.
- Choosing A costs
  - you and player 2: 4 points each
  - player 3 and 4: 8 points each
- If a threshold of players choosing A is reached then
  - you and player 2: get 10 points each
  - player 3 and 4: get 20 points each
- The level of this threshold is changed every 8th round. It is displayed on the screen.
- Each 25 points pays you 1 Euro.

Example

At the beginning of the period you get 8 points. The threshold is 1. Your 3 co-players choose B.

In case you choose A

	you	player 2	player 3	player 4
points at the beginning of the period	8	8	8	8
costs for choosing A	-4	0	0	0
profit for reaching the threshold	+10	+10	+20	+20
period payoff	14	18	28	28

In case you choose B

	you	player 2	player 3	player 4
points at the beginning of the period	8	8	8	8
costs for choosing A	0	0	0	0
profit for reaching the threshold	0	0	0	0
period payoff	8	8	8	8

[9] Experimental Instructions for a small player in treatment 1, translated from German