A Behavioral Theory of Equilibrium Selection

Friedel Bolle

European University Viadrina Frankfurt (Oder)
Department of Business Administration and Economics
Discussion Paper No. 392
January 2017

ISSN 1860 0921

EUROPA-UNIVERSITÄT VIADRINA FRANKFURT (ODER)

A Behavioral Theory of Equilibrium Selection

Friedel Bolle

Europa-Universität Viadrina Frankfurt (Oder)

January 2017

Abstract

Theories about unique equilibrium selection are often rejected in experimental investigations. We drop the idea of selecting a single prominent equilibrium but suggest the coexistence of different beliefs about "appropriate" equilibrium or non-equilibrium play. Our main selection criterion is efficiency applied to all or only to "fair" equilibria. This assumption is applied to 16 Binary Threshold Public Good games where at least *k* of four homogeneous or heterogeneous players have to incur fixed costs in order to produce a public good. The case *k=4* is the Stag Hunt game which is most often used to test equilibrium selection. Our finite mixture model applies with the same parameters (shares of populations, altruism parameters) to the four thresholds *k=1,2,3,4*. The estimated shares of populations are similar in four treatments with identical or different cost/benefit ratios of the players. Our results for *k=4* clearly contradict selection by Risk Dominance and Global Games. In the two (almost) symmetric treatments the Harsanyi/Selten selection explains 40% of the decisions.

JEL codes: C51, C57, C72, D72, H41

Keywords: equilibrium selection, Binary Threshold Public Goods, payoff dominance, risk dominance, Global Games, efficiency, experiment

Friedel Bolle
Europa-Universität Viadrina Frankfurt (Oder)
Grosse Scharrnstrasse 59
D - 15230 Frankfurt (Oder), Germany
bolle@europa-uni.de, Tel.: +49 173 9853143

1. Introduction

The application of game theory is often plagued by the non-uniqueness of equilibria. Prominent examples are coordination games like the Stag Hunt game[1]. There are theoretical attempts to establish a normative theory of equilibrium selection (Harsanyi and Selten, 1988, called HS from now on) but in experiments we rarely observe contribution frequencies which clearly or approximately meet the selected equilibrium or one of the other equilibria. In this paper we want to suggest and test the hypothesis that players behave according to individual beliefs about appropriate equilibria or non-equilibrium strategy profiles, generally called *modes of play*. These modes of play are assumed to be characteristic for certain populations which mainly differ in their beliefs about the priority of efficiency and fairness. We successfully apply our concept to a large variety of Binary Threshold Public Good games which can have more than 30 separate equilibria.

During the last 20 years the discussion about equilibrium selection has more and more turned away from HS. In many experimental investigations equilibria do not play any role at all or are used only as benchmarks. Otherwise, learning to play equilibria (for example, Berninghaus and Ehrhart, 1998) and alternative approaches to equilibrium selection, in particular by Global Games (Carlson and van Damme, 1993) and Quantal Response Equilibria (McKelvey and Palfrey, 1995) have dominated explanations of experimental behavior. In the field of coordination games, the discussion has been focused on the question whether (in games with Pareto-ranked equilibria) payoff-dominance or risk-dominance applies and on the question whether experimental results are close to or converging to Global Games equilibria. Contrary to most other experimental investigations which are concerned with this question we do not investigate 2x2 games but games with four players and two strategies. In the Stag Hunt game, where the risk-dominant and the Global Games predictions (both zero contributions) can be easily computed, our experimental results with about 75% cooperating players in the two (almost) symmetric treatments and more than 90% cooperation in the two asymmetric treatments clearly reject these equilibrium selection principles. HS provide a moderately successful selection for the (almost) symmetric treatments where, according to our estimation, the HS equilibrium is played by 40% of the subjects.

[1] In the Stag Hunt game, players (hunters) can contribute a costly service (go for the stag) or not (succeed in hunting a hare). Hunting the stag is successful only if all players contribute.

Our general hypothesis is that behavior is based on three main, possibly conflicting, requirements[2]:

(i) Consistency (best replies, equilibria)
(ii) Efficiency (social product maximizing strategies)
(iii) Fairness (qualitative or quantitative equality)

The selection of appropriate *modes of play* (selected strategy profiles) takes place on the basis of *individual beliefs* about social norms and the behavior of others. The first principle distinction is whether or not people care about the *consistency* of their beliefs, i.e. whether all strategies should be best replies to the others' strategies (Nash equilibria). Most of our modes of play are equilibrium strategies under certain social or risk preferences but a share of the population may also stick to simple heuristically based modes of play. Efficiency and fairness may be traded off against one another but, because of their simplicity, lexicographic orderings seem to be more sensible to assume[3]. There may be a population of players for which efficiency has priority and another who selects the most efficient among the fair equilibria. There is, however, one further important element of human behavior, namely

(iv) Error.

There are two sources of error, first, concerning the selection of a mode of play and, second, when applying a mode of play. The latter can be generally described by a noise term, in cases of binary decisions by a probability of deviation. The former is problem (game) specific.

In the BTPG games investigated in this paper at least k of n=4 players have to contribute a costly predetermined service in order to produce a public good. Two small populations are assumed not to care about consistency; one is extremely cooperative and the other extremely uncooperative. Players from population P1 (putative pivots, 10-20%) always contribute because they overestimate their own importance or because they are extremely altruistic. Players from population P0 (putative non-pivots, 5%) never contribute because of the opposite reasons. Always or never contributing are equilibria for some k but not for all. Most people, however, are assumed to select only equilibrium

---

2 The principles *efficiency* and *fairness* (equality orientation) are often used for the characterization of experimental results (for example, Engelmann and Strobel, 2004).

3 This is in the spirit of the "Take the Best" heuristic of Behavioral Economics (Gigerenzer, 2008).

modes of play. Some players are assumed to erroneously select the second most efficient equilibrium instead of the most efficient one. Players from the population PE (10-30%) select the most or second most efficient equilibrium; players from the largest population PF (50-65%) select the most or second most efficient of the "fair" equilibria. The reasons of deviations are discussed in more detail in section 3.2. In many games, there are no second best equilibria or error or fairness arguments (in the case of PE) do not apply in which cases no differentiations are made. Note that our definition of populations allows comparing behavior across games. Indeed we find similar or even identical shares of populations for sixteen different BTPG games (four thresholds, four treatments) although the efficient or fair equilibria are rather different across games and treatments. The comparisons of shares rest on successful estimations which means that predicted behavior does not significantly deviate from observed behavior in a χ^2-test. We will always present joint estimations over all four thresholds and also for two of our four treatments.

One apparent objection against our attempt is the question why subjects should stick to "their" selection if they observe others deviating from it. There are three possible reasons for inertia. First, people have identified "the right thing" and stick to it even if others do not. In the finitely repeated Prisoner's Dilemma game often a population P1 of "absolute cooperators" is observed (with a share of 12-13% in Cooper et al., 1996). The second reason might be that subjects observe others not playing "their" mode of play but have no incentive to change behavior. This may be the case for the subpopulation P0 of "absolute defectors". The third argument applies to mixed strategy equilibria and states that deviations from such equilibria are difficult to detect, not even in games with a moderate number of repetitions. In spite of these arguments, in repeated games it remains an empirical question whether people adapt to the play of their co-players or not. If there is adaptation then the application of our static theory can be successful only after the adaptation process has faded out. The question remains why then to investigate repeated games at all. The most important reason is that, in one-shot games, individual mixtures of strategies and population mixtures of pure or mixed strategies cannot be separated.

Our general principles should describe behavior in all games but have to be specified in every application. A first test of our theory considers a class of games with a lot of important applications and with a plethora of equilibria. In our experimental Binary

Threshold Public Good (BTPG) games players i=1,2,3,4 simultaneously contribute (with costs c_i) or not to the production of a public good (benefit G_i) which is produced if there are at least k contributors. The game with k=4 is the Stag Hunt game (Rousseau, 1997, first edition 1762), the game with k=1 is the Volunteer's Dilemma (Diekmann, 1985), and all games can be interpreted as problems with k volunteers necessary or as "Costly Voting games" (Bolle, 2015b, 2016). Problems are naturally framed in a positive frame with the production of a public good ($G_i > c_i > 0$) or in a negative frame with producing a public bad ($G_i < c_i < 0$). Formally, we can transform the two frames into one another and should expect, after applying the transformation, identical behavior. In many other natural examples the question arises whether there is a clear-cut threshold (trigger, tipping point) for a positive or a negative event and whether contributions are binary. In these cases, the BTPG game is, as most 2x2 games are, an approximation which serves the need for simplification for the players as well as the researcher, in both cases because of bounded rationality. For example, Russill and Nyssa (2009) observe a "tipping point trend in climate change communication". For further examples of BTPG games see Bolle (2015).

Let us finally mention a methodological innovation for the estimation of structural models. In economics, practically all structural models are estimated by maximum likelihood which allows comparing the performance of alternative models. The comparison of the best performing model predictions with the data, however, is not offered at all or only graphically. The reason may be that a chi-square test would reject the model. This need not be the case if we use minimum chi-square for the estimation of our model. We will discuss the complementary use of maximum likelihood and minimum chi-square estimations in Section 6.

In the next section we briefly discuss the relevant literature. In Section 3, we introduce BTPG games, derive *equilibrium conditions* and compute equilibria if they are available in closed form. In Section 4, we specify our finite mixture model of (mainly) equilibrium play. Section 5 presents the experiments and provides an overview of the results in terms of average contribution frequencies. In Section 6, our finite mixture model of "equilibrium selection" is tested. Section 7 is the conclusion.

2. Literature

Since Harsanyi and Selten's (1988) suggestion a lot of work has been devoted to the identification of a unique "appropriate" equilibrium. Hansanyi (1995) took a new stance with respect to the priority of payoff-dominance or risk-dominance and Güth and Kalkofen (1989) suggested a related approach. Others favored dynamic concepts (Binmore and Samuelson, 1999) or random deviations (McKelvey and Palfrey, 1995) in order to identify unique equilibria. The suggestion of Carlson and van Damme (1993) to transform common knowledge games into games of incomplete information with private and correlated signals (Global games) has played a major role for equilibrium selection in coordination games. While incomplete information (noise) vanishes play converges, under certain conditions, to one of the pure strategy Nash equilibria of the original game.

To the best of our knowledge, there are only few attempts in the literature to describe behavior as a *finite mixture of equilibrium play* or *best response play* concerning beliefs about the other players. There are models which distinguish types of players with different levels of reasoning (Nagel, 1995; Kübler and Weizsäcker, 2004; Crawford and Iriberri, 2007). Our types, however, do not believe that they are more intelligent or better informed than others. They are distinguished by different beliefs about the appropriate mode of play (mostly equilibrium) for all players. Beliefs (concerning out-of-equilibrium play) are decisive also in dynamic models with incomplete information (McAfee and Schwartz, 1994) but in this literature no attempt is made to analyze the co-existence of different beliefs.

Experimental work on equilibrium selection is often concentrated on the question of which of the two pure strategy equilibria in Stag Hunt games (BTPG games with k=n) and variants of it are played: the payoff-dominant “all contributing” equilibrium or the (mostly) risk-dominant “no one contributing” equilibrium. All studies are with symmetric games, the following also with n=2. Van Huyck et al. (1990) and Rydval und 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 et al. (2011). Whiteman and Scholz (2010), Al-Ubaydli et al. (2013) and Büyükboyacı (2014) investigate the influence of social capital, cognitive ability, own risk aversion, information about others’ risk attitudes, and patience. Spiller and Bolle (2016) investigate the case n=4 with symmetric and asymmetric players who have the same or different cost/benefit ratios and find strong evidence for payoff-dominance. Feltovich and Grossman (2013) investigate the influence of group size (2 to 7 players) and communication on contributions. Without communication, contribution frequencies are about 1/3, independent of group size.

Equilibrium selection is investigated in a meta-study of coordination games with Pareto-ranked equilibria by Blume and Ortmann (2007). They find successful coordination to be the rule rather than the exception. According to Chen and Chen (2011), in a minimum effort game social identity fosters the selection of the most efficient equilibrium. In an experimental investigation of financial attacks (providing a Club Good instead of a Public Good as in BTPG games), Heinemann et al. (2004) find behavior close to the unique Global Game equilibrium. Cabrales et al.'s (2007) experiments show, however, frequent deviations from this equilibrium and emphasize the importance of learning after which behavior can also converge to the payoff-dominant equilibrium. Also Duffy and Ochs (2012) find significant deviations from the Global Game equilibrium.

Experimental studies of BTPG games other than the Stag Hunt game k=n are not concerned with equilibrium selection, in spite (or because) of the tremendous number of equilibria in these games. Experiments with k=1, the Volunteer's Dilemma, are conducted with equal cost/benefit ratios by Diekmann (1985), Franzen (1995), and Goeree et al. (2005). An important result is that, contrary to the theoretical prediction from the unique completely mixed strategy equilibrium, 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 (Fehr and Gächter, 2002), punishment can constitute a Volunteer's Dilemma if 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 (2015a, b) investigate such situations with different costs of the players and find an (incomplete) coordination on the lowest cost player as a volunteer, i.e. there is a tendency towards the asymmetric efficient equilibrium. We will test whether these results can be replicated and extended to higher thresholds. Below, we estimate the share of efficient play in almost symmetric games where coordination on efficient play is difficult and in asymmetric games where it should be easy.

BTPG experiments with intermediate thresholds (in our investigation k=2 or 3 of n=4) have been conducted with k from 2 to 6 and n from 3 to 10, all with at most two different k. For an overview see Spiller and Bolle (2016). 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*.

Spiller and Bolle (2016) investigate the same data set as this paper, however without an attempt to estimate a finite mixture model of equilibrium selection. Their results from non-parametric tests and regression analyses are briefly reported in Section 5.

3. Equilibria of BTPG games and their properties

The general theory of BTPG games is developed in Bolle (2015b). Here we concentrate on results we need for the discussion of our experimental games. In particular we assume players with equal importance for passing the threshold. In the *positive frame*, there is a set of n players $N = \{1, \dots, n\}$ who can contribute (with costs $c_i > 0$) or not (without costs) to the production of a public good. If a certain threshold k of contributions is surpassed, the public good is produced and the players earn $G_i > c_i$. If a player does not contribute and the project is not launched his revenue is 0. 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 (k=n) are symmetric. With different cost/benefit ratios also mixed strategy equilibria are asymmetric but they may be viewed as "less asymmetric" and "more fair" than asymmetric pure strategy equilibria.

The case $G_i < c_i < 0$ is called the *negative frame*. In the following sense, it is the "mirror image" of the positive 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.

Let us assume that the players' contribution probabilities are $p = (p_i)_{i=1,\dots,n}$. $Q = Q(p)$ denotes the probability of success, i.e. that k or more players contribute to the production of the public good. Q_{-i} (Q_{+i}) denote the probability of success if i does not contribute (contributes). These probabilities depend 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

(1) $$R_i(p) = G_i * Q(p) - p_i c_i$$
$$= G_i * Q_{-i} + p_i * [G_i * q_i - c_i] .$$

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

(2) $$\partial R_i / \partial p_i = G_i * q_i - c_i = 0.$$

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.

(3) $$R_i = G_i * Q_{-i} = G_i * Q_{+i} - c_i.$$

Proposition 1: The following statements apply in equilibrium:

(i) If i plays a strictly mixed strategy, then $q_i = r_i = c_i / G_i$.

(ii) $q_i > r_i$ implies $p_i = 1$ and $q_i < r_i$ implies $p_i = 0$.

(iii) $R_i = G_i Q_{-i}$ applies for $p_i < 1$ and $R_i = G_i Q_{+i} - c_i$ for $p_i > 0$.

Proof: (1), (2) and (3).

The case $\boldsymbol{k = n}$

This case is called the **Stag Hunt game,** first discussed by Rousseau (1997, first edition 1762). There are two symmetric pure strategy equilibria, namely $p = (0, \dots 0)$, $p = (1, \dots, 1)$ and, possibly, a completely mixed strategy equilibrium which is derived from (2) and $q_i = \prod_{j \neq i} p_j$. It follows $p_i = \left(\prod_j r_j\right)^{1/(n-1)} / r_i$. The condition of the existence of this equilibrium is $p_i \leq 1$ for all i. This condition is always fulfilled for n=2 or if all r_i are identical. Smaller r_i are connected with larger p_i. Because of (3) and $Q_{-i} = 0$ the mixed strategy equilibrium yields zero revenues. There are possibly also pure/mixed strategy equilibria where some players contribute with probability 1 and the others play the mixed strategy equilibrium of a reduced Stag Hunt game. According to Proposition 1, those who contribute with probability 1 earn $R_i = G_i * Q_{+i} - c_i \geq 0$ (if $R_i < 0$, this isn't an equilibrium) and the mixed strategy players earn zero. Because of Proposition 1 (iii), $p = (1, \dots, 1)$ is the payoff-dominant equilibrium.

Let us, for this case and certain parameters, determine also the Global Game equilibrium and the risk dominant equilibrium under the definition of Harsanyi and Selten (1988).

Unfortunately, the application of Global Games is hindered by the possible dependency of the resulting (in many cases unique) equilibrium on the distribution of "noise" (Frankel et al, 2003) and by a lack of methods to compute the equilibrium in cases of asymmetric games with more than 2 players. Therefore, based on results by Frankel et al. (2003) we determine the global game equilibrium only for k=4 (Stag Hunt game) which, for our experimental games, coincides with the Risk Dominant equilibrium according to the definition of Harsanyi and Selten (1988).

Proposition 2: $a^* = (0, \ldots, 0)$ is the unique global game equilibrium of a BTPG game with k=n.

Proof: Appendix A1.

Frankel et al. (2003) require actions to be strategic complements. This requirement is not fulfilled in cases k<n. Others increasing their contribution (in our case from 0 to 1) can make it advantageous for i to reduce his contribution (from 1 to 0). On the first glance this is a bit surprising because Frankel et al.'s (2003) theory can be applied to the quite similar case of financial attacks against a currency (Heinemann et al., 2004). If the model assumes binary choices and if the attack is successful when at least k<n players join the attack then this is not a BTPG game. Because only players joining the attack can profit the players provide a Club Good and not a Public Good. For Club Goods actions are strategic complements.

Let us now turn to Risk Dominance as defined by Harsanyi and Selten (1988).

Proposition 3: In the case k=n, if $r_i > \prod_{j \neq i} r_j$ for all i then $(0, \ldots, 0)$ risk dominates all other equilibria.

Proof: Appendix A1.

Corollary: In our four experimental treatments with cost/benefit ratios of r_i=0.4 in the two almost symmetric treatments and (0.225, 0.25, 0.275, 0.3) and (0.1, 0.2, 0.3, 0.4) in the asymmetric treatments, the risk dominant equilibrium in the games with k=4 is $(0, \ldots, 0)$.

The case $k = 1$

This case is called the **Volunteer's Dilemma**, first investigated by Diekmann (1985, 1993). There are n pure strategy equilibria where exactly one player contributes. The only completely mixed strategy equilibrium is derived from (2) and $q_i = \prod_{j \neq i}(1 - p_j)$. It follows $p_i = 1 - \left(\prod_j r_j\right)^{1/(n-1)}/r_i$. Therefore this equilibrium exists under the same conditions as that of the Stag Hunt game. Smaller r_i are connected with smaller p_i (regarded as counterintuitive by Diekmann, 1993). Because of Proposition 1 and $Q_{+i} = 1$, in this equilibrium players earn $R_i = G_i - c_i$, i.e. as much as players who always contribute.

The case 1<k<n

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, 2015b) and π is derived from

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

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 (iii)). If the r_i are unequal then the system of equations (2) has to be solved with q_i being a more complicated function of p_i than (4).

The number of equilibria

A completely mixed strategy equilibrium depends only on $r_i = c_i/G_i$ and therefore applies in the positive ($G_i > c_i > 0$) as well as in the negative ($G_i < c_i < 0$) frame. Pure strategy equilibria and equilibrium selection, however, correspond only after applying the strategically neutral transformation.

If 1<k<n, n>3, then completely mixed strategy equilibria can be determined only by numerical methods. For n=4, four polynomial equations of degree 3 with four variables have up to 12 different solutions, though not necessarily real numbers and not necessarily in $(0,1)^4$. For our experimental case n=4 and if $c_i/G_i = r_i$ are not equal we find numerically (with a lot of parameter variations) mostly up to two completely mixed strategy equilibria, in rare cases also more than two.

Independent of whether or not r_i are equal, there are many more pure/mixed strategy equilibria (see Table 1). In the case k=1, the Volunteer's Dilemma, there are four pure strategy equilibria where exactly one player contributes, there is possibly one completely mixed strategy equilibrium (see above), there are up to four additional equilibria where one player plays p_i=0 and the others according to the completely mixed strategy equilibrium of the Volunteer's Dilemma with n=3, and there are up to six equilibria where two players play p_i=0 and the other two according to the completely mixed strategy equilibrium of the Volunteer's Dilemma with n=2. The number of equilibria for k=2 and k=3 are derived accordingly. In the Stag Hunt game, however, no pure/mixed strategy equilibrium exists where the player with the highest r_i contributes with p_i=1. This results from the others contributing with such probabilities that their expected revenue is 0.

Threshold k	1	2	3	4
# pure str. equ.	4	7	5	2
# compl. mixed equ.	≤ 1	≤ 2*	≤ 2*	≤ 1
# pure/mixed equ.	≤10	≤ 24	≤ 24	≤ 6

Table 1: Number of equilibria in the positive frame if the threshold is "k contributions from n=4 players".

Explanatory remarks: * For the parameters estimated below there are exactly two equilibria. Computations with many different parameter constellations resulted often in less than two completely mixed strategy equilibria and in rare cases in more than two.

The HS selection for games with identical $\boldsymbol{r_i}$

In the case of symmetric games, Harsany and Selten (1992) restrict their selection to the set of symmetric equilibria. These can generically be ordered according to Pareto-dominance. For BTPG games we extend the HS definition of symmetry to games with identical $r_i = c_i/G_i$.

Proposition 2: In a BTPG game with identical r_i 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 $p_i = \pi''(\pi')$ otherwise $p_i = 0$ (1).

Proof: Appendix A1.

4. The selection of modes of play in BTPG games

As already outlined in the introduction we allow the coexistence of different beliefs about the appropriate selection of a *mode of play*. Our first distinction is between equilibrium and non-equilibrium players. While equilibrium requires a certain consistency of beliefs, we assume also two (small) populations of non-equilibrium players who do not care about the beliefs of others. There are absolute cooperators (population P1) who believe that their own contribution is decisive with a high probability ($q_i > r_i$ in Proposition 1) and who therefore always contribute ($p_i=1$) in the positive frame. There is also a population P0 with opposite beliefs ($q_i < r_i$) that never contributes ($p_i=0$) in the positive frame. P1 may consist of extremely altruistic players and P0 of extremely spiteful players. Close to the P0 players are free riders who may be characterized by optimistic maximax strategies (resulting in $p_i=0$ for games with $k<4$ and $p_i=1$ for $k=4$ in the positive frame) and risk averters who may be characterized by pessimistic maximin strategies (resulting in $p_i=0$ for games with $k>1$ and $p_i=1$ for $k=1$ in the positive frame). We introduce P0 players as counterparts of P1 players but it may turn out that, in other applications, the alternatives maximax or minimax strategies are more different and more successful.

Equilibrium players believe that the appropriate mode of play is defined as the most efficient among the Nash equilibria (population PE) or among the "fair" Nash equilibria (population PF). An *efficient* mode of play maximizes the social product (the sum of incomes). "*Fairness*" is used here mainly in the sense of *"equality"* and, as a concession to bounded rationality, it is defined only qualitatively with the binary values "equal" and "unequal". Modes of play p=(1,1,1,1) or p=(0,0,0,0) or a completely mixed strategy with $0<p_i<1$ for all i are considered as (qualitatively) equal and therefore fair; all other modes of play p are defined as unequal and unfair. Of course, fairness has many other facets which may be important in other applications and even below we will discuss also another interpretation of fairness but our main assumption remains that mixed strategy equilibria are fairer than asymmetric pure strategy equilibria.

The assumption that all players behave according to these four classes is too strong, however. We assume that most efficient equilibria are selected only by the subpopulations PE1 and PF1 and introduce two further classes PE2 and PF2 where "most efficient" is substituted by "second most efficient". Such behaviors may be assumed to be errors or they may indicate deviations from lexicographic preferences. PE2 players may be concerned "a bit" about fairness and PF2 players "a bit" about error of others and risk. In a game with the threshold k=2 and different costs $c_1<c_2<c_3<c_4$, the most efficient Nash equilibrium is p=(1,1,0,0). Although in PE efficiency has priority over fairness, some dissatisfaction of the contributing players can be expected. In particular, player 2 may ask whether player 3 should not also contribute, at least sometimes. In one-shot games[4], however, they cannot coordinate on alternatingly playing (1,1,0,0) and (1,0,1,0). Instead of that we assume the subpopulation PE2 playing according to the unique mixed strategy equilibrium $(1,v_2,v_3,0)$ where v_i denotes the mixed strategy equilibrium probabilities of the Volunteer's Dilemma with n=2. In games with k=1 and k=3, PE2 people are assumed to play according to the unique strategies $(v_1,v_2,0,0)$ and $(1,1,v_3,v_4)$; in the game with k=4 dissatisfaction arguments are without bite and we assume both subpopulations playing (1,1,1,1). In the population PF a fair equilibrium is selected but players may be concerned about riskiness in a loose and weak sense. PF1 players select (1,1,1,1) in the game with k=4, and the (if it exists) unique completely mixed strategy equilibrium M_1 for k=1, and the most efficient H_k of completely mixed strategy equilibria (if existent) in the games with k=2 and k=3. Population PF2 trades off efficiency against equilibria with less frequent contributions (on average less risky concerning errors or deviant beliefs of their co-players) and select the (if existent) unique completely mixed strategy M_4 for k=4, and the second most efficient completely mixed strategy equilibrium L_k for k=2 and k=3. For k=1, the completely mixed equilibrium is the only fair equilibrium. In the error interpretation of PF2's behavior, the selection of L in the games with k=2 and k=3 is the consequence of ignorance of H_k, in particular if we assume equilibria not to be computed but approximately known from a lifelong experience with similar situations.

Table 2 provides an overview about populations and their selections of modes of play in the case with different costs, i.e. in our treatments A and B. For the thresholds k=1 and k=4, some modes of play of different populations coincide. In particular, the Stag Hunt game alone cannot provide an estimation of the hypothetical populations.

Asymmetry is generic but many example games are symmetric. This poses a problem for efficiency play in our almost symmetric games for cases k=1 and k=3 where there are two efficient asymmetric equilibria and where it seems to be impossible to coordinate actions for playing one of these. Our treatments S+ and S- have identical r_i=0.4 and, in

[4] In our experiments repeated games with a stranger design were played, i.e. for every repetition the four players in a game were randomly composed.

S+, c1=c2=4, c3=c4=8. This allows to define equilibria E1 for k=2 and k=4 and equilibria E2 for k=1 and k=3 as indicated in Table 2. Our new definition of E1 is the play of (v_1,v_2,0,0), (1,1,0,0), (1,v_2,v_3,0), (1,1,1,1) for k=1,2,3,4. Population E2 in treatments S+ (correspondingly in S-) is assumed to be concerned also with fairness and they may argue as follows: Players 3 and 4 have not only larger costs but also larger benefits; if player 3 contributes and player 1 not and if the project is launched then player 3 earns 12 and player 1 earns 8. Therefore population E2 favors switching the roles of small and large players compared with E1's selection of equilibria.

Pop	Characterization	Modes of play for k=1,2,3,4
PE1	Most efficient equilibrium	(1,0,0,0), (1,1,0,0),(1,1,1,0), (1,1,1,1)
PE2	"Second" most efficient equilibrium	(v_1, v_2,0,0), (1, v_2, v_3,0),(1,1, v_3, v_4), (1,1,1,1)
PF1	Most efficient fair equilibrium	M_1, H_2, H_3, (1,1,1,1)
PF2	Second most efficient fair equilibrium	M_1, L_2, L_3, M_4
P1	Putative pivots	(1,1,1,1) for all k
P0	Putative non-pivots	(0,0,0,0) for all k

Table 2: *Subpopulations for games with* $c_1<c_2<c_3<c_4$

Explanatory remarks: M_1 (M_4) denotes the (if existent) unique completely mixed strategy equilibrium in the game with k=1 (k=4). v_2 denotes the symmetric equilibrium contribution probabilities of the Volunteer's Dilemma with two players. H denotes the most efficient of the (mostly two) completely mixed strategy equilibria with k=2 and k=3, L the second most efficient one.

The shares of the six subpopulations constitute the first five parameters of our finite mixture model. The shares are assumed to be independent of the threshold k and the player type (c_i, G_i). The pure strategy modes of play are, except in extreme cases, independent of social preferences. Mixed strategy equilibria, however, vary with the social preferences which constitute one (in the two almost symmetric treatments) or four (in the two asymmetric treatments) additional parameters.

Social preferences are introduced as *altruism and/or warm glow* in the spirit of Andreoni's (1989, 1990) suggestion. They change the game only insofar as the cost/benefit ratios 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

(5) $R_i = Q * (G_i + a_i * G_{-i}) - (1 - b_i) * p_i * c_i.$

This results in the equivalent to (2),

(6) $(G_i + a_i * G_{-i}) * q_i - (1 - b_i) * c_i = 0.$

Proposition 4: The introduction of altruistic and/or warm glow players results in an equilibrium condition for mixed strategies

(7) $q_i = r_i * s_i$ with $s_i = \frac{G_i(1-b_i)}{G_i + a_i * G_{-i}}$.

Proof: (6).

For the sake of simplicity we assume $a_i = 0$ so that $s_i = 1 - b_i$ does not depend on c_i or G_i. In the following we will assume that players with equal r_i have equal s_i, while otherwise the s_i may be different. This is in the spirit of role dependent preferences (Bolle and Otto, 2016) which assume that social preferences are adapted in an evolutionary process in order to improve the strategic position of a player. For an investigation concerning the evolutionary stability[5] of altruistic preferences see Bester and Güth (1998) and Bolle (2000). Roles are defined by the strategic situation of a player, here the cost/benefit ratios of all players. In the almost symmetric treatments one $s_i = s$ is estimated and in the asymmetric treatments there are four (one for every player). These values are independent of the populations and the thresholds.

The last parameter of our model describes an average random deviation (perturbation) probability from the strategies selected. Small and large contribution probabilities are thus moved to the middle as hypothesized in Prospect Theory when probability weighting functions are introduced. The perturbation probability should be small and is indeed estimated as smaller than 3.3% for four of the six separately estimated data sets.

4. Experiments and overview of results

All our experimental games are with four players. In Treatment S+ (almost symmetric, positive frame), players 1 and 2 with $(c_i, G_i) = (4,10)$ Lab-Dollars are called small players; players 3 and 4 with $(c_i, G_i) = (8,20)$ are called large players. In Treatment S- (almost symmetric, negative frame) G_i and c_i have the same absolute values as in S+ but are both negative, i.e., players earn a profit by contributing and suffer a loss if the threshold is

[5] Evolutionary stable preferences depend on the parameters of the game (Heifetz, 2007) and thus such an approach challenges the stability of preferences. Note, however, that also other approaches as many variants of Prospect Theory (Kahnemann and Tversky, 1979) do this. Bolle and Otto (2016) comment on the plausible extent of role dependent variability.

surpassed. Again, players 1 and 2 are called small players and 3 and 4 large players. All players have a cost/benefit ratio c_i/G_i =0.4.

In the asymmetric treatments A and B benefits were G_i=20 and costs varied. In Treatment A, contribution costs (c_i) were (4.5, 5, 5.5, 6) and cost/benefit ratios (r_i)=(0.225, 0.25, 0.275, 0.3) had a small spread. In Treatment B costs were (2, 4, 6, 8) and cost/benefit ratios (0.1, 0.2, 0.3, 0.4) showed a large spread. The costs and benefits of a player define his *type*. A player kept his type during the whole experiment. Every subject participated in only one treatment.

Treat-ment	Endow-ment	costs c_i	Benefits G_i	c_i/G_i	#sessions (at V, at TU)
S+	8	(4,4,8,8)	(10,10,20,20)	0.4	(10, -)
S-	20	(-4,-4,-8,-8)	(-10,-10,-20,-20)	0.4	(10, -)
A	8	(4.5, 5, 5.5, 6)	20	(0.225, 0.25, 0.275, 0.3)	(6, 12)
B	8	(2, 4, 6, 8)	20	(0.1, 0.2, 0.3, 0.4)	(10, 6)

Table 3: Game parameters (in lab dollars) in the four treatments for players i=1,2,3,4 and number of sessions with eight subjects either at TU (Technische Universität Berlin) or V (Europa-Universität Viadrina Frankfurt (Oder)).

We conducted the experiments as computerized laboratory experiments (implemented in a z-tree program design, Fischbacher, 2007) at two locations, the Vialab (V) of the Europa-Universität Viadrina in Frankfurt (Oder) and in the experimental laboratory of the Technische Universität (TU) Berlin. Table 4 describes the experimental parameters and how many sessions of a treatment were conducted at TU and Viadrina.

A session consisted of 32 games with the same eight subjects. In every session there were four (in treatments S+ and S-) or two (in treatments A and B) players of each type. In each of the 32 periods they were allocated randomly to two experimental groups under the restriction that in every group two (in treatments S+ and S-) or one player of each type was present. So there are 36 different groups in treatments S+ and S- and 16 different groups in treatments A and B. In each session every threshold k=1, 2, 3, 4 was played in eight periods in a row. During 32 periods all thresholds were adopted in a random order but with the restriction that, in the 10 sessions with treatments S+, S-, and B at the Viadrina, each *k* was played either 2 or 3 times at each of the four positions. In the 12 sessions with treatment A at TU each k was played three times at each of the four positions, in the six sessions of A at Viadrina and B at TU, each k was played once or twice in each position.

Subjects were not informed about the order of the thresholds in the beginning, but only when the threshold changed. We mentioned already that 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.

Before subjects played the games, they were given printed instructions and had the possibility to ask 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 until 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 eighth 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.

In all of the 32 periods players were endowed with 8 Lab-Dollars (treatments S+, A, B) or 20 Lab-dollars (treatment S-). If the threshold of k contributions was reached or surpassed, all players received the benefit G_i (suffered losses in treatment S-); otherwise they received nothing. Their total income in a period consisted of their endowment minus their costs of contributing (if they contributed) plus benefits (if the threshold was reached or surpassed). One Lab Dollar was worth 4 Eurocents. Participants earned between 17 and 36 Euros with an average of 28.11 Euros. Sessions lasted roughly 45 minutes.

Average contribution probabilities

In Tables A2 and A3 (Appendix A3, adopted from Spiller and Bolle, 2016), *average contribution frequencies* $ACF(k, c_i, G_i)$ are reported for different thresholds k and different player types described by (c_i, G_i). Non-parametric tests are carried out based on session averages. The stylized conclusions for *treatments S+ and S-* are:

- *Small and large players show similar ACFs except for k=4, i.e. there is little evidence of efficient play with small players contributing more frequently than large players.*

- *ACFs in the positive and the negative frame are mirrored*, i.e. $ACF(k, c_i, G_i) = 1 - ACF(5 - k, -c_i, -G_i)$.

Stylized conclusions for treatments A and B are:

- There are no systematic differences between V and TU subjects in treatment B; there are some differences in treatment A.

- There is a tendency towards efficiency. Generally, lower cost players contribute with higher probability. In particular, the k least cost players contribute more than the other players.

For all treatments we find:

- ACFs increase with the threshold.

- The predictions of Global Games and Risk Dominance (Propositions 2 and 3 for k=4), namely that no one will contribute in the game with k=4, are clearly rejected.[6]

Further results are reported in Spiller and Bolle (2016) where also regression analyses are carried out which show traces of dynamics, in particular a trend towards more cooperation in later periods of S+ and S-.

5. A finite mixture model of equilibrium selection

In treatments A and B we found hints that efficient play may have a certain influence on average contribution frequencies. On the other hand, it is clear that, except in the Stag Hunt game (k=4), only a small part of the population can have played efficient pure strategies. The question is whether efficiency players exist at all or whether there is only a general tendency that players with smaller c_i contribute with higher probability. Therefore we now turn to individual contributions.

We call a player's number of contributions to the public good in the eight repetitions of decisions (for a certain threshold k) the *individual contribution frequency ICF*. The distributions of these ICFs are provided in Appendix A5 for treatments, subject pools,

[6] Note that, in treatment S-, k=1 is the Stag Hunt game and no one contributing is the cooperative equilibrium.

player types, and thresholds, for treatments S+ and S- also for the first half of experiments (period<17) and the second half (period>16). As an example, Figure 1 provides the ICFs of treatment S+, aggregated over small and large players. There are some players who choose the same action eight times (ICF=0 or ICF=8), but this number varies over games and player types. As we assume the shares of the subpopulations to be independent of the threshold k and the player type, we conclude that the share of P1 cannot be large and that the share of P0 must be even smaller. Also efficiency players from PE cannot be frequent. So we expect the bulk of subjects to belong to the PF subpopulations.

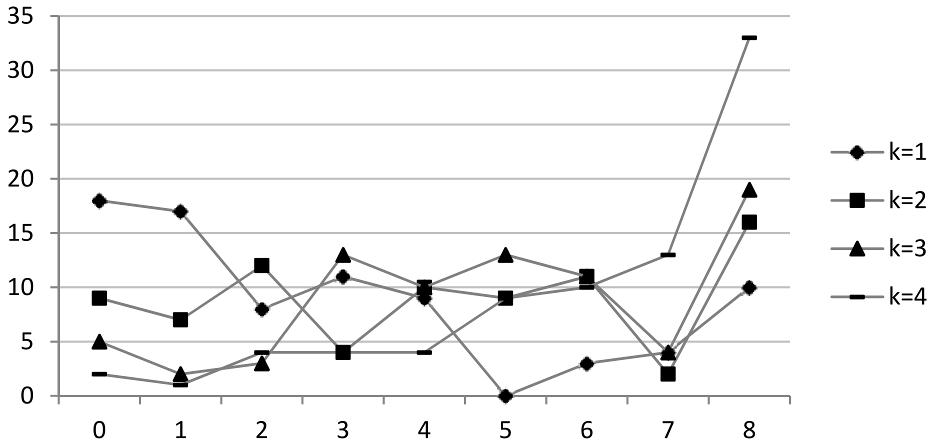

Figure 1: Frequency distribution of individual contribution frequencies (ICFs) in treatment S+. k= threshold. For every k, 8 decisions by 80 individuals.

Methodological issues

For the estimation of our model we need a hypothesis about the variation of membership in the six subpopulations (Table 2) during the 32 periods of an experiment. A regression model for single decisions predicts each decision *separately*. In a finite mixture model this would mean:

HypNo: Players may switch randomly between the six subpopulations in every of their 32 decisions. The shares of the subpopulations remain constant, however.

Alternative hypotheses are:

HypThresh: Players keep their membership of one of the six subpopulations during the eight decisions under a certain threshold. Between games with different thresholds they may change their membership. The shares of the subpopulations remain constant however.

HypAll: Players keep their membership of one of the six subpopulations during the 32 decisions of a session.

HypNo almost obliterates the assumption of consistently acting subpopulations. Its consequence is that players in a certain game (defined by the threshold k) act according to an average contribution probability, i.e. overall we should observe a binomial distribution of ICFs for a certain k. This is apparently not the case in Figure 1. *HypThresh* results in a mixture of binomial distributions of ICFs and *HypAll* in a mixture of the product of four binomial distributions. *HypAll* is possibly too demanding. A new game with a new threshold requires a new evaluation and, similar as perturbations occur in the play of strategies, they may also occur in the choice of the appropriate mode of play. In Appendix A3 we compute the likelihood functions for the three hypotheses and the chi-square function for *HypThresh*. We present also a table with the results of Maximum Likelihood estimations under the three hypotheses. As *HypThresh* is the clear winner of this competition we employ it in the following.

In the following, S+/S-(per<17) denotes the aggregated data from the two treatments S+ and S- in the first half of the session (period<17), i.e., the first two experiments of a session. A_{TU} denotes the data from treatment A at the TU laboratory. Etc.. The estimation of our model is carried out separately for each of the six data sets S+/S-(per<17), S+/S-(per>16), A_{TU}, A_{V}, B_{TU}, and B_{V} and jointly for the data of the same treatments A_{TU}+A_{V}, B_{TU}+B_{V}, and S+/S-(per<17) + S+/S-(per>16). The separate investigation of the early (per<17) and late (per>16) decisions in S+/S- is motivated by the detection of dynamics in a regression analysis (Spiller and Bolle, 2016). The separate investigation of TU and V data is due to the four significant differences in treatment A (Table A4). For the estimation of the model parameters we have employed the Maximum Likelihood method as well as the Minimum Chi-square method (Berkson, 1980; Newey and West, 1987). Maximum Likelihood allows to identify significant differences between alternative models and to evaluate the aggregation of data sets; the chi-square score is a measure of the absolute fit between model and data. Results of these estimations are presented in Tables 4 and 5.

Although the parameter estimations by Minimum Chi-square and Maximum Likelihood are asymptotically equivalent, real estimations differ to a certain degree. The loglikelihood scores are improved by 3-13 points when we employ Maximum Likelihood instead of Minimum Chi-square (see Table 4) while the Chi-square values increase considerably and signal significant deviations between the predictions from maximum likelihood estimations of parameters and results. This gives rise to the question whether Minimum Chi-square leads to over-fitting, i.e. the question whether it can compensate structural weaknesses of the model. We think this is unlikely because most model variations as, for example, neglecting the mostly smallest subpopulation, P0 with a share of 3-6%, lead to strongly increasing χ^2 scores. We take the stance that the loss of degrees of freedom by estimating the seven or ten parameters is best taken into account with corresponding tests and estimation procedures. Likelihood ratio tests and applications of the Akaike and Bayes information criteria AIC and BIC should be carried out on the basis of Maximum Likelihood estimations; a Chi-square test of the model fit on the basis of Minimum Chi-square estimations. Of course, we should keep in mind that the loss of degrees of freedom by the number of estimated parameters is generally true only for linear models (Andrae et al., 2010), and that it is unclear when numbers are large enough for asymptotic properties to apply. The fine differentiation of ICFs allows testing a particularly detailed structure but it leads to many small frequencies. As a rule of thumb, the estimated frequencies in Chi-square tests should not be smaller than 5. This requirement is not fulfilled in our estimations. Note, however, that the danger under such circumstances is to produce too large χ^2 scores, i.e. there is an increased danger of rejecting an adequate model. Therefore corrections as the *Williams correction*[7] decrease the computed χ^2 scores. In Table 7 we report uncorrected χ^2 scores. Because of our large sample size, the Williams correction factor for χ^2 is smaller than 1.02 and would not change our evaluations considerably. We carry out, however, an additional estimation based on categories where ICFs {0,1,2}, {3,4,5}, and {6,7,8} are aggregated. As a result, the $p(\chi^2)$ values are either about the same or considerably larger than in the disaggregated model.

[7] χ^2 is divided by the Williams correction factor which is always larger than 1. In our investigation it is maximal in A_V (1.0191). χ^2 is then reduced from 141.9 to 139.2 and p increases from 0.066 to 0.088.

The fit of the finite mixture model

The chi-square scores and the average log-likelihood scores indicate that our model shows the worst fit in the almost symmetric cases S+/S- and the best fit in the case of the highly asymmetric treatment B and for the subject pool TU. For all data sets except S+/S- (per<17) the chi-square scores indicate a sufficient fit (not rejected on the 5% level). In the maximum likelihood estimations of treatments S+/S- the score of the joint maximum likelihood estimation of 1329.5 is 33.5 points worse than the separate estimations for per<17 and per>16, 700.8+595.2=1296.0. For treatment A, the score of the joint estimation A_{TU}+A_V, 980.0, is worse by 27.9 points compared to the score in the separate estimations. According to a Likelihood Ratio test the separate estimation is significantly better ($p<10^{-7}$ in the case of treatment A, even smaller in S+/S-), in spite of the additional 10 parameters (7 parameters in S+/S). Also AIC and BIC favor the separate estimation. The separate estimations in treatment B are worse by only 10.7 points and are justified according to the Likelihood ratio test with p=0.018 and (just) AIC but not according to BIC. The minimum chi-square estimation confirms the first two comparisons, but it allows merging the subject pools in treatment B.

		Minimum χ^2			Minimum χ^2_r		Maximum Likelihood			
Data	N	χ^2	$p(\chi^2)$	-logL	χ^2_r	$p(\chi^2_r)$	χ^2	$p(\chi^2)$	-logL	-logL/N
S+/S- per<17	320	171.0	0.002	712.1	24.7	0.479	216.4	$<10^{-6}$	700.8	2.190
S+/S- per>16	320	146.1	0.060	602.9	38.6	0.040	174.2	0.001	595.2	1.860
S+/S- all	640	190.8	$<10^{-4}$	1342.5	22.1	0.683	248.8	$<10^{-9}$	1329.5	2.077
A_{TU}	384	121.0	0.405	610.5	24.4	0.328	134.5	0.142	604.5	1.574
A_V	192	141.9	0.066	350.9	24.8	0.304	177.3	0.003	347.6	1.810
A_{TU}+ A_V	576	181.7	10^{-4}	986.7	32.4	0.070	208.5	$<10^{-6}$	980.0	1.701
B_{TU}	192	124.2	0.300	291.0	18.6	0.667	368.3	0	279.2	1.454
B_V	320	122.0	0.382	549.3	20.6	0.546	143.4	0.056	544.4	1.701
B_{TU}+ B_V	512	135.5	0.129	841.3	24.4	0.328	162.6	0.004	834.3	1.629

Table 4: Minimum Chi-square and Maximum likelihood estimation of the finite mixture model with six data sets under *HypThresh*.

Explanatory remarks: χ^2 is determined from the estimation of 144 cells of 16 mixtures of binomial distributions (2 treatments times 2 player types times four games in S+/S-, 4 player types times 4 games in treatments A and B) with nine different ICFs; therefore df=144-16-10=118 for $p(\chi^2)$ from treatments A and B and df= 121 for treatments S+/S-. For χ^2_r only three classes of ICFs are defined; therefore df=48-16-10=22 for $p(\chi 2_r)$ from treatments A and B and df= 25 for treatments S+/S-.

	ε	α_1	α_{E1}	α_{E2}	α_{F1}	α_{F2}	α_0	s_1	s_2	s_3	s_4
S+/S-	0.114	0.168	0	0.008	0.223	0.398	0.202	-	-	-	0.708
per<17	0.010	0.025	0.010	0.026	0.029	0.026	-	-	-	-	(0.039)
S+/S-	0.021	0.122	0.035	0.109	0.406	0.259	0.070	-	-	-	0.814
per>16	(0.005)	(0.017)	(0.020)	(0.023)	(0.027)	(0.022)	-	-	-	-	(0.024)
A_{TU}	0.020	0.180	0.083	0.214	0.322	0.171	0.029	1.414	1.880	1.686	1.546
	0.002	0.020	0.022	0.023	0.025	0.019	-	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$
A_V	0.063	0.149	0	0.104	0.498	0.186	0.062	1.894	1.779	1.642	1.485
	0.008	0.025	0.039	0.032	0.024	0.026	-	$<10^{-3}$	$<10^{-3}$	$<10^{-3}$	$<10^{-3}$
B_{TU}	0.022	0.226	0.172	0.132	0.208	0.213	0.050	3.270	1.378	1.072	0.784
	0.004	0.027	0.035	0.0405	0.033	0.026	-	0.234	0.098	0.070	0.055
B_V	0.032	0.110	0.119	0.217	0.321	0.188	0.045	3.294	1.331	1.020	0.757
	0.005	0.019	0.029	0.034	0.032	0.019	-	0.157	0.075	0.058	0.043

Table 5: Parameters estimated by minimum chi-square, standard errors in parentheses. Explanatory remarks. Standard errors are estimated by the square roots of the diagonal elements of the inversion of the Hessian. A peculiarity of the standard errors is their tiny values for the s_i parameters in A_{TU} and A_V. This results from the fact that the estimated parameters "just" allow the existence of completely mixed strategies.

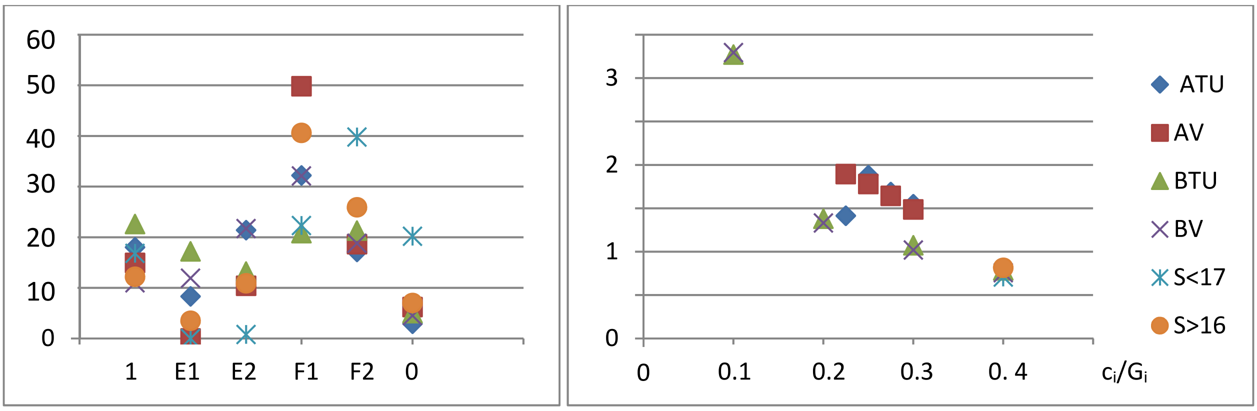

Figure 2: Shares (%) of subpopulations (left) and altruism/warm glow parameters of player types characterized by c_i/G_i. S<17 stands for S+/S-(per<17).

The parameters of the estimations according to minimum chi-square are shown in Table 5 and Figure 2. Note that $\alpha_0 = 1 - \alpha_1 - \alpha_{E1} - \alpha_{E2} - \alpha_{F1} - \alpha_{F2}$. The corresponding contribution probabilities for S+/S-(per>16) and the two larger data sets A_{TU} and B_V of the asymmetric treatments are presented in Appendix A4. First, we observe that the perturbation probability ε is small and has its largest values for the data set with the worst fit, S+/S-(per<17). Otherwise ε is at most 0.063 and, on average, 0.030; therefore it certainly does not dominate the intrinsic structure of the model.

Second, we find some considerable differences between S+/S-(per<17) on the one hand and S+/S-(per>16) as well as the other data sets on the other. The estimated parameters for S+/S-(per<17) are either outliers or at least extreme compared with the other estimations. From the first to the second half of the experiments with S+ and S- subjects learn to contribute: the shares of efficient and F1 play increased and those of P0 and F2 decreased. Leaving aside the outliers from S+/S-(per<17), the shares of populations P1, PE2, PF2, and P0 are rather similar while the shares of PE1 and PF1 are more variable. For four of the six data sets, PF1 is the largest population; for treatments S+ and S- PF1 is the HS selection.

Third, the altruism/warm glow parameters vary little between subject pools but a lot between treatments. Figure 2 shows a strong negative correlation between $r_i = c_i/G_i$ and s_i which may be expressed by a linear or hyperbolic function, in the latter case $r_i * s_i \approx 0.35$. If our model is correct the estimation of different preferences across player types but not across populations means that preferences are not stable. Our explanation is that preferences are "role dependent" where a role is narrowly defined as a player in a certain game. But it does not make sense to define thousands of different roles; therefore the same role should be taken also in "similar" strategic situations. Preferences guide behavior and are therefore similar to commitments[8] which allow players to gain higher *material* profits. Whether and under which conditions evolutionary stable preferences are described by $r_i * s_i = const$ may be investigated as in Bester and Güth (1998); but such a theoretical investigation is beyond the scope of this paper.

Roles in bargaining are discussed by Bolle and Otto (2016). Envy towards one's bargaining partner generally improves the bargaining results of a player (except when both show so much spite that no agreement can be reached) but although, except for $r_i = 0.4$, all the estimated s_i are larger than 1 and thus indicate spite or cold prickle, in BTPG games things are more difficult. In the mixed strategy equilibrium of the Volunteer's Dilemma a player's increasing altruism improves his material success, in the

[8] That's the point in strategic delegation (Vickers, 1985; Fershtman et al.,1991).

mixed strategy equilibrium of the Stag Hunt game it reduces his material success. The same contrary effects are valid when comparing the completely mixed strategy equilibria with k=2 and k=3.

In general, our equilibrium selection hypothesis has turned out to be successful, but of course minor adaptations to other applications may be necessary. An example is the substitution of the small population P0 by a population of maximax players (free riders) or maximin players (risk averters). In our data set, in both cases only one of the four games (k=1 or k=4) is affected. For many simple games, the predictions for several populations coincide. Therefore the complete model should be re-tested preferably in a rich environment and not for single 2x2 games. On the other hand, the theory must be applicable also to such "simple games".

7. Conclusion

The main message from this investigation is that Nash equilibria can explain behavior but that, first, people have individual beliefs about the appropriate equilibrium, second, that people have adopted "role dependent" social preferences, and, third, that there is a certain level of random and perhaps also systematic error. With some qualifications, behavior in our four treatments can be explained by a finite mixture model with six populations who are guided by different principles for the selection of (mostly equilibrium) modes of play. About 80% of the subjects either play most efficient equilibria or most efficient fair equilibria. Some fuzziness is introduced by a perturbation probability (about 3%) and by reducing the "most efficient" requirement to a "second best" level.

The almost symmetric treatments S+ and S- are estimated jointly, i.e., with the same parameters. Contrary to many linear Public Good experiments (e.g. Dufwenberg et al., 2011) no effect of framing a decision positively or negatively is observed. Comparing early and late decisions in treatments S+ and S- shows that there is a trend towards more cooperation. The consequence is that only decisions from the second half of the experiment fit our static model with a non-significant chi-square score. In the moderately asymmetric treatment A and the considerably asymmetric treatment B a strong (A) and a weak (B) subject pool effect is observed. Only in the latter case is the estimation of the model with the joint data from two laboratories at different universities successful. Parameters are always estimated jointly for games with thresholds k=1,2,3,4. Across treatments, the shares of populations do not differ considerably. The altruism parameters do differ but they are rather similar for players with the same cost/benefit ratios. For increasing cost/benefit ratios, altruism increases (s_i decreases). Figure 2 suggests a close relationship between the cost/benefit relation of a player and his altruism/ warm glow parameter. We interpret this as an indication for role dependent preferences.

People often doubt that game theoretic equilibria have any meaning for behavior. In the Volunteer's dilemma not even qualitative predictions seem to apply. (a) In the completely mixed strategy equilibrium of the symmetric Volunteer's Dilemma the probability of producing the public good should decrease with the number of players n, but it is shown to increase (Diekmann, 1986). (b) In the completely mixed strategy equilibrium of the asymmetric Volunteer's Dilemma, the probabilities of contribution should increase with the cost/benefit ratios of players, but average observed frequencies are shown to decrease (Diekmann, 1993). However, (b) is implied in our finite mixture model because of the existence of efficiency players and because altruism parameters counteract the influence of cost/benefit ratios. (a) follows from the existence of the group P1 whose members always contribute. The larger the number of players n, the larger the probability that a member of P1 is present. In many other investigations of BTPG games not even an attempt is made to match observed behavior with equilibria (see Section 2).

The other extreme game, the Stag Hunt game, is the favorite example for discussing coordination problems. In our experiments the "optimistic" subpopulations P1, PE1, PE2, and PH who have an aggregate share of about 75% choose to always contribute. A share of about 20% PL people contribute with a probability of 2/3 and only about 5% P0 people do not contribute at all. This result clearly contradicts the selection of P0 by risk dominance or Global Games. It poses the question how to explain behavior in some experiments with the Stag Hunt game (van Huyck et al., 1990; Rydval und Ortmann, 2005) with intermediate contribution probabilities or convergence to zero contributions. The usual result for the wider class of games with Pareto-ranked equilibria is, however, successful coordination (Blume and Ortmann, 2007).

Dynamics *need not* be strong, in particular in a stranger design. For repetitions in a partner design, however, a behavioral drift may be the dominant phenomenon. The consequence would be that subpopulations could no longer be assumed to be constant. From the perspective of our model the most interesting question is whether even after many repetitions several modes of play survive or whether convergence to a single equilibrium is the rule.

Acknowledgement: I am grateful for the funding by the Deutsche Forschungsgemeinschaft (project BO 747/14-1)

References

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

Andrae, R., Schulze-Hartung, T., Melchior, P. (2010). Dos and don'ts of reduced chi-squared. *arXiv preprint arXiv:1012.3754*.

Andreoni, J. (1989) “Giving with impure altruism: applications to charity and Ricardian equivalence” *The Journal of Political Economy*, 1447-1458.

Andreoni, J. (1990) “Impure altruism and donations to public goods: a theory of warm-glow giving” *The Economic Journal*, 464-477.

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

Battalio, R., Samuelson, L., Van Huyck, J. (2001) “Optimization incentives and coordination failure in laboratory stag hunt games” *Econometrica*, 69(3), 749-764.

Berkson, J. (1980) “Minimum Chi-Square, not Maximum Likelihood!” *The Annals of Statistics* 8 (3), 457-487.

Berninghaus, S. K., & Ehrhart, K. M. (1998). Time horizon and equilibrium selection in tacit coordination games: Experimental results. *Journal of Economic Behavior & Organization*, *37*(2), 231-248.

Bester, H. and Güth, W., 1998. Is altruism evolutionarily stable. Journal of Economic Behavior & Organization 34, pp. 193–209.

Binmore, K., & Samuelson, L. (1999). Evolutionary drift and equilibrium selection. *The Review of Economic Studies*, *66*(2), 363-393.

Blume, A., & Ortmann, A. (2007). The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria. *Journal of Economic Theory*, *132*(1), 274-290.

Bolle, F. (2000): "Is Altruism Evolutionarily Stable? And Envy and Malevolence? Remarkson Bester and Güth", Journal of Economic Behavior and Organization 42, 131-133.

Bolle, F. (2015a). A Note on Payoff Equivalence of the Volunteer's Dilemma and the Stag Hunt Game and Inferiority of Intermediate Thresholds. *International Game Theory Review*, *17*(03), 1550004.

Bolle, F. (2015b) “Costly Voting - A General Theory of Binary Threshold Public Goods” *Discussion Paper, Frankfurt (Oder)*.

Bolle, F., Otto, P.E. (2016). "Role-dependent social preferences" Economica, DOI: 10.1111/ecca.12180.

Bullinger, A. F., Wyman, E., Melis, A. P., 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.

Cabrales, A., Nagel, R., Armenter, R. (2007). Equilibrium selection through incomplete information in coordination games: an experimental study. *Experimental Economics*, *10*(3), 221-234.

Carlsson, H., Van Damme, E. (1993). Global games and equilibrium selection. *Econometrica*, 989-1018.

Chen, R., Chen, Y. (2011). The potential of social identity for equilibrium selection. *The American Economic Review*, 2562-2589.

Chen, X. P., Au, W. T., Komorita, S. 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.

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1996). Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, *12*(2), 187-218.

Crawford, V. P., Iriberri, N. (2007). Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?. *Econometrica*, *75*(6), 1721-1770.

Diekmann, A. (1985) “Volunteer’s Dilemma” *Journal of Conflict Resolution*, Vol. 29/4, 605-610.

Diekmann, A. (1986). Volunteer’s Dilemma. A Social Trap without a Dominant Strategy and some Empirical Results. In *Paradoxical Effects of Social Behavior* (pp. 187-197). Physica-Verlag HD.

Diekmann, A. (1993) “Cooperation in an Asymmetric Volunteer’s Dilemma Game – Theory and Experimental Evidence” *International Journal of Game Theory* 22, 75-85

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

Diekmann, A., & Przepiorka, W. (2015b). “Take One for the Team!” Individual Heterogeneity and the Emergence of Latent Norms in a Volunteer's Dilemma. *Social Forces*, sov107.

Downs, A. (1957) “An Economic Theory of Democracy” *New York: Harper*.

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

Duffy, J., & Ochs, J. (2012). Equilibrium selection in static and dynamic entry games. *Games and Economic Behavior*, *76*(1), 97-116.

Engelmann, D., and Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *The American Economic Review*, *94*(4), 857-869.

Erev, I., Rapoport, A. (1990) "Provision of Step-Level Public Goods: The Sequential Contribution Mechanism" *Journal of Conflict Resolution*, Vol. 34/3, 401-425.

Fehr, E., Gächter, S. (2002) "Altruistic punishment in humans" *Nature*, 415(6868), 137-140.

Feltovich, N., & Grossman, P. J. (2013) "The effect of group size and cheap talk in the multi-player stag hunt: experimental evidence" *(No. 52-13). Monash University, Department of Economics*.

Fershtman, C., Judd, K. L., & Kalai, E. (1991). Observable contracts: Strategic delegation and cooperation. *International Economic Review*, 551-559.

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

Frankel, D. M., Morris, S., & Pauzner, A. (2003). Equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory*, *108*(1), 1-44.

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

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on psychological science*, *3*(1), 20-29.

Goeree, J.K., Holt, C.,Moore, A. (2005) "An Experimental Examination of the Volunteer's Dilemma", *www.people.virginia.edu/~akm9a/Volunteers Dilemma/vg_paperAug1.pdf.*

Güth, W., Kalkofen, B. (1989), "Unique Solutions for Strategic Games - Equilibrium Selection Based on Resistance Avoidance". Lecture Notes in Economics and Mathematical Systems, Vol. 328.

Harsanyi, J. C., Selten, R. (1988) A general theory of equilibrium selection in games. *MIT Press Books*.

Harsanyi, J.C. (1995) "A New Theory of Equilibrium Selection for Games with Complete Information", *Games and Econonmic Behavior* 8, 91–122.

Heifetz, A., Shannon, C., Spiegel, Y. (2007): "What to maximize if you must", Journal of Economic Theory 133, 31-57.

Heinemann, F., Nagel, R., & Ockenfels, P. (2004). The theory of global games on test: Experimental analysis of coordination games with public and private information. *Econometrica*, *72*(5), 1583-1599.

Kahneman, D., Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica:, 263-291.

Kübler, D., Weizsäcker, G. (2004). “Limited depth of reasoning and failure of cascade formation in the laboratory“, The Review of Economic Studies, 71, 425-441.

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

McKelvey, R. D., Palfrey, T. R. (1995). “Quantal response equilibria for normal form games”. Games and Economic Behavior, 10(1), 6-38.

Nagel, R. (1995) "Unraveling in Guessing Games: An Experimental Study". The American Economic Review, Vol. 85.

Newey, W.K., West, K.D. (1987) "Hypothesis Testing with Efficient Method of Moments Estimation” *International Economic Review* 28 (3), 777-787.

Przepiorka, W., 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.

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

Russill, C., and Nyssa, Z. (2009). The tipping point trend in climate change communication. *Global environmental change*, *19*(3), 336-344.

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

Spiller, J., Bolle, F., (2016). “Experimental investigations of binary threshold public good games”. Manuscript, Europa-Universität Viadrina Frankfurt (Oder).

Van Huyck, J. B., Battalio, R. C., Beil, R. O. (1990) “Tacit coordination games, strategic uncertainty, and coordination failure” *American Economic Review*, 80(1), 234-248.

Vickers, J. (1985). Delegation and the Theory of the Firm. *The Economic Journal*, *95*, 138-147.

Whiteman, M. A., Scholz, J. T. (2010) “Social Capital in Coordination Experiments: Risk, Trust and Position”. Working Paper, Southern Illinois University Carbondale.

Appendix

A1 Proofs

Lemma 1 (for the proof of Proposition 2): In a BTPG game with k=n, the strategy profile $a^* = (0, \dots, 0)$ is a local potential maximizer.

Proof: Let $a = (a_i, a_{-i})$ denote a vector of actions of the players with $a_i = 1$ denoting contribution to the production of the public good and $a_i = 0$ denoting non-contribution. Adapting the requirements on a local potential maximizer $a^* = (0, \dots, 0)$ from Frankel et al. (2003) to our binary choice game we have to find a function $v(a) = v(a_i, a_{-i})$ which takes a strict local maximum at $a = a^*$ and find positive numbers μ_i so that

$$v(0, a_{-i}) - v(1, a_{-i}) \leq \mu_i[R_i(0, a_{-i}) - R_i(1, a_{-i})] = \begin{cases} \mu_i c_i & \text{for } a_{-i} \neq (1, \dots, 1) \\ -\mu_i(G_i - c_i) & \text{for } a_{-i} = (1, \dots, 1) \end{cases}$$

for all i. The function $v(a) = \sum_i R_i(a)$ and $\mu_i = 1$ fulfill these conditions.

Proposition 2: $a^* = (0, \dots, 0)$ is the unique global game equilibrium of a BTPG game with k=n.

Proof: Because of Frankel et al. (2003), Theorem 1, the equilibrium is unique, because of Theorem 2 it is a pure strategy Nash equilibrium, and because of Theorem 3 (which needs Lemma 1 above) it is independent of the distribution of noise.

Proposition 3: In the case k=n, if $r_i > \prod_{j \neq i} r_j$ for all i then $(0, \dots, 0)$ risk dominates all other equilibria p.

Proof: According to Harsanyi and Selten (1988), for the question whether a mixed or pure strategy equilibrium p risk dominates another equilibrium p' first the *bicentric prior* of p and p' is derived. For BTPG games we have to determine, for every $0 \leq t \leq 1$, whether $a_i = 1$ or $a_i = 0$ is a best reply of player i to the other players contributing with probabilities $t * p_{-i} + (1 - t) * p'_{-i}$. The shares of t values with $a_i = 1$ constitute a vector x of prior probabilities. With these priors the *tracing procedure* is carried out where for every $0 \leq t \leq 1$ equilibria are determined in a game where player i assumes that, with probability t, the BTPG game is played and with 1-t the other players decide according to the prior probability. If there is a continuous path of equilibria from t=0 to t=1 then the corresponding equilibrium for t= 1 is selected.

The bicentric priors of the equilibria $(1,\ldots,1)$ and $(0,\ldots,0)$ are $x_i^* = r_i$ and are at least as large as the bicentric priors of any strategy profile p and $(0,\ldots,0)$. Because of $q_i = \prod_{j\neq i} r_j$ the best reply to these priors is $p_i = 0$ (Proposition 1 (ii)). Then there is a constant path of equilibria $(0,\ldots,0)$ for all t which constitutes the generically unique risk dominant equilibrium. (Lemma 4.17.7 in Harsanyi and Selten, 1988).

Proposition 2: In a BTPG game with identical r_i 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 $p_i = \pi''(\pi')$ 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.

A2 Average contribution frequencies

	pos. frame S+		neg. frame S-	
k	Small	Large	Small	L arge
1	0.35*	0.37*	0.30	0.26
2	0.49*	0.56	0.43*	0.39*
3	0.61*	0.63*	0.57*	0.49*
4	0.74§	0.81	0.75§	0.59

Table A1: Average contribution probabilities (ACPs) in treatments S+ and S-. *Source: Spiller and Bolle (2016).*

Explanatory note: Small player type S with $(G_S,c_S)=(10,4)$ and large player type L with $(G_L,c_L)=(20,8)$ in the positive frame. k= threshold. HS= equilibrium according to Harsanyi and Selten (1992) as described in Proposition 3 and without altruism or warm glow. § Two-sided Wilcoxon matched pairs-test for small vs. large players. * One-sided Wilcoxon matched pairs-test of non-increasing ACPs for k (position of *) 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 on averages from 10 sessions and p<0.05.

	Exp A (A_{TU} + A_V)				Exp B (B_{TU} + B_V)			
c_i/G_i k	0.225	0.25	0.275	0.3	0.1	0.2	0.3	0.4
1	0.389	**0.497***	0.333*	0.250*	0.676	0.344*	0.227*	0.277*
2	0.622	0.625	**0.483***	0.483*	0.781	0.613	0.398*	0.418*
3	0.733	0.792	0.733*§	**0.559*§**	0.930	0.840	0.688*§	0.637*§
4	0.997	0.948	0.931	**0.944***	0.984	0.945	0.918	0.883*

Table A2: Average contribution frequencies in treatments A and B. *Source: Spiller and Bolle (2016).*

Explanatory note: There are four significant differences (bold types) between V and TU subjects in two-sided Wilcoxon tests on the 5% level, in three cases higher probabilities in TU, in one case in V. All differences between threshold k and threshold k+1 are significant in two-sided Wilcoxon match-pairs tests (except in three cases) on the 5% level. * (§) Significant differences between player types compared with type c_i/G_i =0.225 (0.25) in treatment A and c_i/G_i =0.1 (0.2) in treatment B in a two-sided Wilcoxon test on the 5% level. Tests are based on averages from 16 sessions (A) and 18 sessions (B) and p<0.05.

A3 Chi-square and Log-likelihood functions

In the following description the treatment and the subject pool are kept constant.

For the three alternative hypotheses about the duration of subpopulation (group) membership we use the following representation of our data. Let k denote the threshold of a game, h the player type, and i=ICF the individual contribution frequency of a player in the 8 repetitions of a game. $F(k,h,i)$ denotes how many players h in a game with threshold k contributed i times to the production of the public good. (See A3.) $H(h,i_1,i_2,i_3,i_4)$ denotes the number of players h who contributed i_1 times in the game with threshold k=1, i_2 times in the game with threshold k=2, etc.. Most of the $4*9^4$=26244 cells of H are empty, i.e. H(.)=0.

The theoretical prediction about frequency distributions depend on the hypothesis about the duration of membership in a subpopulation. In Section 5.3 we regarded the three alternatives *HypNo, HypThresh, and HypAll.*

In the subpopulation $j \in \{1, E1, E1, F1, F2, 0\}$ a player h in game k contributes with probability p_j*(k,h/X) which depends on the parameter vector X=(s_1, s_2, s_3, s_4). The shares of the groups are α_1, α_{E1}, α_{E2}, α_{F1}, α_{F2} and $\alpha_0=1-\alpha_1-\alpha_{E1}-\alpha_{E2}-\alpha_{F1}-\alpha_{F2}$. The parameters are defined in Section 5.3 and are assumed to be independent of the group. The equilibrium contribution probabilities p_j*(k,h/X) and their determination are described in Section 3.2. Taking into account the perturbation probability ε, the effective contribution probability is $p_j(k,h/X,\varepsilon) = (1-\varepsilon)\, p_j^* + \varepsilon(1- p_j^*)$. This perturbation is most important for the pure strategies p_j^*=1 or 0.

Under *HypNo* the probability of player h contributing in a game with threshold k is

$$prob_{No}(\mathrm{k,h/X,\varepsilon}) = \sum_{j=1,E1,E2,F1,F2,0} \alpha_j * \mathrm{p_j(k,h/X,\varepsilon)}$$

and the log-likelihood function is

$$logL_{No} = \sum_{k,h,i} F(k,h,i) * (\mathrm{i} * \log(prob_{No}(\mathrm{k,h/X,\varepsilon})) + (8-i) * \log(1 - prob_{No}(\mathrm{k,h/X,\varepsilon})).$$

Under HypThresh, after 8 independent repetitions the number of contributions i=ICF is binomially distributed with a probability B(i,8; p_j(k,h/X)). Taking into account the different groups j we get a probability for i contributions by player h in a game with threshold k of

$$prob_{Thresh}(k,h,i/X) = \sum_{j=1,E1,E2,F1,F2,0} \alpha_j * \mathrm{B(i,8;\ p_j(k,h/X,\varepsilon))}.$$

The log-likelihood function is

$$logL_{Thresh} = \sum_{k,h,i} F(k,h,i) * \log(prob_{Thresh}(\mathrm{k,h/X,\varepsilon})).$$

Under *HypAll*, the probability of observing (i_1,i_2,i_3,i_4) in the games with k=1, 2, 3, 4 is

$$prob_{All}(h,i_1,i_2,i_3,i_4/X,\varepsilon) = \sum_{j=1,E1,E2,F1,F2,0} \alpha_j * \prod_{\mathrm{k}} \mathrm{B(i_k,8;\ p_j(k,h/X,\varepsilon))}.$$

The log-likelihood function is

$$logL_{All} = \sum_{h,i_1,i_2,i_3,i_4} H(h, i_1, i_2, i_3, i_4) \ * \log(prob_{All}(h, i_1, i_2, i_3, i_4/X, \varepsilon)).$$

The chi-square score is computed only for *HypThresh*:

$$\chi^2 = \sum_{k,h,i} (F(k, h, i) - \mathrm{N} * \ prob_{Thresh}(\mathrm{k, h/X, \varepsilon}))^2/(N * prob(\mathrm{k, h/X, \varepsilon}))$$

with $N = \sum_i F(k, h, i)$= number of players of type h in game k (independent of h and k in our experiment).

X and the shares of groups and ε are chosen such that -logL or χ^2 is minimized. For this purpose the Nelder-Mead algorithm from the R-library is used with a large number of starting values. For the model specification outlined in subsections 3.2 and 5.2 HypThresh is the clear winner of the competition as Table A1 shows. Note that computations of S+/S-(per<17) or S+/S-(per>16) for $-logL_{All}$ requires the computation of products of two binomial distributions for different combinations of thresholds k.

Data	$-logL_{No}$	$-logL_{Thresh}$	$-logL_{All}$
S+/S-(all per.)	3282.4	1329.5	1463.3
A_{TU}	1529.2	604.5	705.9
A_V	869.1	347.6	384.1
B_{TU}	733.9	279.2	373.9
B_V	1282.6	544.4	599.7

Table A3: Maximum likelihood scores of the finite mixture model under the hypotheses *HypNo, HypThresh*, and *HypAll*. For treatments A and B, ten parameters are estimated, for treatments S+/S-, six.

A4 Equilibrium contribution probabilities with Parameters from Table 5

	share	k=1	k=2	k=3	k=4
P1	.122	(1, 1)*	(1, 1)*	(1, 1)*	(1, 1)
PE1	.035	(.67, 0)	(1, 0)	(1, 0.67)	(1, 1)
PE2	.109	(0, .67)	(0, 1)	(0.67, 1)	(1, 1)
PF1	.406	(.31, .31)	(.56, .56)	(.85, .85)	(1, 1)
PF2	.259	(.31, .31)	(.15, .15)	(.44, .44)	(.69, .69)
P0	.070	(0, 0)*	(0, 0)	(0, 0)	(0, 0)

Table A4 : Share of groups and equilibrium probabilities of player types (small players, large players) for treatment S+. Parameters estimated from data S+/S-(per>16). * Non-equilibria.

	share	k=1	k=2	k=3	k=4
P1	.180	(1, 1, 1, 1)*	(1, 1, 1, 1)*	(1, 1 ,1, 1)*	(1, 1, 1, 1)
PE1	.083	(1, 0 ,0, 0)	(1, 1, 0, 0)	(1, 1, 1, 0)	(1, 1, 1, 1)
PE2	.214	(.53,.68,0,0)	(1,.54,.53,0)	(1,1,.54,.54)	(1, 1, 1, 1)
PF1	.322	($0^{\$}$, .32,.31, .31)	(.08,.61,.55,.55)	(.55,.84,$1^{§}$,.77)	(1,1,1,1)
PF2	.171	($0^{\$}$,.32,.31,.31)	(.45, .16, $0^{§}$,.23)	(.91,.39,.45,.45)	($1^{\$}$, .68,.69,.69)
P0	.029	(0,0,0,0)*	(0,0,0,0)	(0,0,0,0)	(0,0,0,0)

Table A5 : Share of groups and equilibrium probabilities of player types with cost/benefit ratios (.225, .25, .275, .3). Parameters estimated from data A_{TU}. * Non-equilibria. § More exactly: \$ $0<p_1<10^{-8}$ or $1-10^{-8}<p_1<1$. § $0<p_3<10^{-4}$ or $1-10^{-4}<p_1<1$.

	share	k=1	k=2	k=3	k=4
P1	.110	(1,1,1,1)*	(1,1,1,1)*	(1,1,1,1)*	(1,1,1,1)
PE1	.119	(1,0,0,0)	(1,1,0,0)	(1,1,1,0)	(1,1,1,1)
PE2	.0217	(.73,.67,0,0)	(1,.69,.73,0,0)	(1,1,.70,.69)	(1,1,1,1)
PF1	.321	(.40,.23,.35, .34)	(.69,.46,.61,.60)	(.93, .80, .88, .87)	(1,1,1,1)
PF2	.188	(.40,.23,.35, .34)	(.07,.20,.12,.13)	(.31,.54,.39,.40)	(.60, .77,.65, .66)
P0	.045	(0,0,0,0)*	(0,0,0,0)	(0,0,0,0)	(0,0,0,0)

Table A6 : Share of groups and equilibrium contribution probabilities of player types with cost/benefit ratios (.1, .2, .3, .4). Parameters estimated from data B_V. * Non-equilibria.

A5 Data: The distribution of individual contribution frequencies (number of players with a certain ICF)

	Periods 1- 16, small players									Periods 17 – 32, small players								
k ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1	5	4	2	3	1	0	2	1	2	6	5	0	2	3	0	0	0	4
2	2	1	3	1	2	2	1	1	3	3	2	4	1	2	3	4	1	4
3	1	1	1	3	5	4	3	1	5	3	1	1	4	1	1	0	0	5
4	2	1	3	0	4	0	3	4	3	0	0	0	0	0	2	2	3	13
	Periods 1- 16, large players									Periods 17 – 32, large players								
k ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1	5	2	2	3	4	0	1	2	1	2	6	4	3	1	0	0	1	3
2	0	2	3	2	2	2	2	0	3	4	2	2	0	4	2	4	0	6
3	1	0	0	6	3	5	3	1	5	0	0	1	0	1	3	5	2	4
4	0	0	1	4	0	6	2	4	3	0	0	0	0	0	1	3	2	14

Table A7: Frequency distribution of ICFs (individual contribution frequencies) in Treatment 1. k= threshold. For every k, 8 decisions by 40 individuals.

	Periods 1- 16, small players									Periods 17 – 32, small players								
k ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1	3	1	0	3	3	4	1	4	1	16	3	1	0	0	0	0	0	0
2	2	4	2	2	3	4	1	1	1	3	2	4	3	1	4	1	1	1
3	1	1	4	2	5	1	1	3	2	0	2	2	3	5	3	0	1	4
4	0	0	0	0	5	4	7	4	0	2	1	0	0	0	4	5	2	6
	Periods 1- 16, large players									Periods 17 – 32, large players								
k ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1	3	1	1	3	5	1	1	2	3	14	4	1	1	0	0	0	0	0
2	4	2	3	2	1	4	1	1	2	4	3	3	2	1	2	3	1	1
3	3	0	3	2	1	2	5	2	2	4	2	1	2	1	3	1	3	3
4	2	1	1	1	3	2	4	1	5	3	0	2	1	0	2	1	5	6

Table A8: Frequency distribution of ICFs (individual contribution frequencies) in Treatment 2. k= threshold. For every k, 8 decisions by 40 individuals.

		Player type 1: r1 =0.225									Player type 2: r2 =0.25								
k	ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1		9	0	2	2	3	2	0	2	4	5	2	3	4	2	2	0	1	5
2		3	0	2	1	2	3	1	1	11	2	1	4	2	3	3	0	1	8
3		1	1	1	2	0	3	1	1	14	0	1	0	1	1	0	6	4	11
4		0	0	0	0	0	0	0	0	24	0	0	0	0	0	1	1	1	21
		Player type 3: r3 =0.275									Player type 4: r4 =0.3								
k	ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1		8	5	1	1	2	1	2	0	4	8	4	4	3	3	0	0	1	1
2		5	1	1	4	1	4	2	1	5	4	2	1	5	2	3	1	2	4
3		0	1	1	2	2	2	2	1	13	2	2	2	3	0	4	2	3	6
4		1	0	0	0	0	0	1	1	21	0	0	0	0	0	0	1	1	22

Table A9: Frequency distribution of ICF. Data A_{TU} (Treatment A, subject pool TU). For every game and every player type 24 players (=number of ICF=sum of rows).

		Player type 1: r1 =0.225									Player type 2: r2 =0.25								
k	ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1		1	4	0	4	1	0	1	0	1	0	2	1	1	1	2	1	1	3
2		0	2	3	1	1	2	0	1	2	0	0	3	1	1	0	0	2	5
3		0	1	0	2	2	2	1	1	3	0	1	0	1	1	3	1	1	4
4		0	0	0	0	0	0	0	1	11	0	0	0	1	0	1	0	1	9
		Player type 3: r3 =0.275									Player type 4: r4 =0.3								
k	ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1		3	0	3	2	3	1	0	0	0	1	4	5	1	0	0	0	1	0
2		0	1	2	5	2	1	0	0	1	1	2	2	1	0	2	3	0	1
3		1	0	0	1	3	1	2	3	1	0	1	4	1	2	2	0	1	1
4		0	1	0	0	0	0	1	0	10	0	1	0	1	0	0	0	1	9

Table A10: Frequency distribution of ICF. Data A_V (Treatment A, subject pool V). For every game and every player type 12 players (=number of ICF=sum of rows).

		Player type 1: r1 =0.225									Player type 2: r2 =0.25								
k	ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1		1	0	1	2	1	0	0	1	6	4	1	3	3	0	0	0	0	1
2		0	0	0	1	0	1	2	1	7	2	0	0	0	3	1	0	2	4
3		0	0	0	0	1	0	0	0	11	0	0	0	0	0	1	1	3	7
4		0	0	0	0	0	0	0	1	11	0	0	0	0	0	0	1	1	10
		Player type 3: r3 =0.275									Player type 4: r4 =0.3								
k	ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1		6	2	0	1	1	0	0	2	0	3	3	2	0	1	0	0	0	3
2		5	1	1	0	1	1	1	0	2	4	0	0	2	1	0	0	0	5
3		1	0	0	3	1	1	0	1	5	3	0	0	0	0	4	1	1	3
4		0	0	0	0	1	0	0	1	10	1	0	1	0	0	1	0	1	8

Table A11: Frequency distribution of ICF. Data B_{TU} (Treatment B, subject pool TU). For every game and every player type 12 players (=number of ICF=sum of rows).

	Player type 1: r1 =0.225									Player type 2: r2 =0.25								
k ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1	2	0	1	0	2	4	5	2	4	5	2	3	4	0	1	0	1	4
2	0	0	1	1	2	3	4	7	2	2	0	4	2	1	2	2	2	5
3	0	0	0	0	1	1	2	3	13	0	0	1	2	0	2	4	3	8
4	0	0	0	0	0	0	0	3	17	0	1	0	0	0	0	0	4	15
	Player type 3: r3 =0.275									Player type 4: r4 =0.3								
k ICF	0	1	2	3	4	5	6	7	8	0	1	2	3	4	5	6	7	8
1	10	2	2	3	1	0		0	2	9	2	2	2	4	0	0	0	1
2	4	1	4	1	4	1	2	1	2	6	2	3	2	2	1	0	2	2
3	0	1	3	1	1	3	2	2	7	1	2	0	1	3	1	3	6	3
4	0	0	0	1	1	1	1	2	14	0	0	0	0	0	2	2	2	14

Table A12: Frequency distribution of ICF. Data B_V (Treatment B, subject pool V). For every game and every player type 20 players (=number of ICF=sum of rows).

A6. Instructions

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 for Treatment A

- 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. [In other instructions: Player 2, 3, or 4]
- 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**
  - o **you and player 2** **4 points each**
  - o **player 3 and 4** **8 points each**
- If a **threshold of players choosing A** is reached then
  - o **you and player 2** **get 10 points each**
  - o **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