CESifo Working Papers
www.cesifo.org/wp

Minimum Participation Rules for the Provision of Public Goods

Anke Gerber
Jakob Neitzel
Philipp C. Wichardt

CESIFO WORKING PAPER NO. 3733
CATEGORY 13: BEHAVIOURAL ECONOMICS
FEBRUARY 2012

CESifo *Working Paper No. 3733*

Minimum Participation Rules for the Provision of Public Goods

Abstract

This paper considers the endogenous formation of an institution to provide a public good. If the institution governs only its members, players have an incentive to free ride on the institution formation of others and the social dilemma is simply shifted to a higher level. Addressing this second-order social dilemma, we study the effectiveness of three different minimum participation requirements: 1. full participation / unanimity rule; 2. partial participation; 3. unanimity first and in case of failure partial participation. While unanimity is most effective once established, one might suspect that a weaker minimum participation rule is preferable in practice as it might facilitate the formation of the institution. The data of our laboratory experiment do not support this latter view, though. In fact, weakening the participation requirement does not increase the number of implemented institutions. Thus, we conclude that the most effective participation requirement is the unanimity rule which leaves no room for free riding on either level of the social dilemma.

JEL-Code: C720, C920, H410, D020.

Keywords: public goods, coalition formation, endogenous institutions.

Anke Gerber
University of Hamburg
Department of Economics
Von-Melle-Park 5
Germany – 20146 Hamburg
anke.gerber@wiso.uni-hamburg.de

Jakob Neitzel
University of Hamburg
Department of Economics
Von-Melle-Park 5
Germany – 20146 Hamburg
jakob.neitzel@wiso.uni-hamburg.de

Philipp C. Wichardt
Department of Economics
University of Bonn and LMU Munich
Kaulbachstraße 45
Germany – 80539 Munich
philipp.wichardt@uni-bonn.de

February 6, 2012
We are grateful to Hans Carlsson, Fabian Herweg, Håkan J. Holm, Mathias Kifmann, Charles Kolstad, Andreas Lange, Sandra Ludwig, Axel Ockenfels, Bettina Rockenbach, Klaus Schmidt, Miriam Schütte, Gari Walkowitz, Achim Wambach as well as seminar participants at the University of Zurich and the University of Cologne for helpful comments and discussions.

1 Introduction

The provision of public goods is a social dilemma that has attracted a lot of attention ever since the seminal article of Samuelson (1954). Many solutions have been proposed so far, all relying on some institution that sets the rule of the game such as to provide individuals the incentive to contribute the efficient amount to the public good. These institutions can be characterized either by centralized or decentralized sanctioning (punishment and reward) and the sanctioning can either be formal (monetary transfers) or informal (e.g. social ostracism).[1]

While institutions have been shown to be effective not only theoretically but also empirically, the main hindrance to their adoption is the fact that the implementation itself is a public goods problem. Obviously, any selfish individual prefers others to install the institution and provide the public good. Hence, it is questionable whether the institution will be implemented at all if membership is voluntary. In the present paper, we focus on this second order public goods problem and study the endogenous formation of an institution that, once established, enforces efficient contributions to a public good by all its members.

There is a large theoretical literature on endogenous institution (coalition) formation in the context of global environmental problems (see e.g. Carraro and Siniscalco, 1993; Barrett, 1994; Carraro and Marchiori, 2003; Finus and Rundshagen, 2003) yielding the pessimistic result that stable coalitions usually comprise only few members and hence efficiency is low.[2] In these models there are no restrictions on the size of coalitions. A straightforward question therefore is whether efficiency could be increased under appropriate restrictions on coalition size. A natural restriction would be a minimum participation requirement and in fact minimum participation rules are very common in international environmental agreements. The Kyoto protocol, for example, had to be ratified by at least

[1]Examples for centralized formal sanctioning institutions are the mechanisms proposed by Groves and Ledyard (1977), Moore and Repullo (1988), Abreu and Sen (1990), Palfrey and Srivastava (1991), Jackson (1992), Falkinger (1996) and more recently Gerber and Wichardt (2009a, 2009b). Decentralized sanctioning with punishment or rewards being executed by players themselves have been studied among others by Fehr and Gächter (2000, 2002), Masclet et al. (2003) and Sefton et al. (2007).

[2]Stability here refers to the notion of cartel stability introduced by d'Aspremont et al. (1993).

55 parties accounting for at least 55 percent of greenhouse gas emissions in 1990 before entering into force. Other treaties like the Convention for the Protection of the Marine Environment of the North-East Atlantic required ratification by all abutting nations. Rutz (2001) has analyzed data provided by the International Center for Earth Science Information Network (CIESIN) showing that the number of international environmental treaties that do not contain any minimum participation requirement is negligible.

Focusing on the consequences of minimum participation requirements, this paper addresses the empirical question which minimum participation rule is most effective in terms of maximizing overall efficiency in the context of a simple linear public goods game. While a stricter requirement leads to the implementation of larger institutions, and, hence to an increase in efficiency whenever an institution is actually implemented, a weaker requirement may be more effective overall since it renders the implementation of an institution less vulnerable to a coordination failure.

In the experiment, groups of four players interacted in an institution formation game under different minimum participation rules covered in three different treatments: In treatment 1 an institution was only implemented when all four group member joined the institution (unanimity rule). In treatment 2 at least three group members had to join the institution and in treatment 3 the minimum participation requirement was relaxed to three in a second institution formation stage whenever institution formation failed under the unanimity rule in the first stage.

According to the data, overall efficiency turns out to be largest in those treatments that restrict to or start with the unanimity rule (treatments 1 and 3). What is more, the total number of institutions is not larger under a weaker minimum participation rule than under the unanimity rule.[3] From an applied point of view, the results of our experiment, thus, show that weakening participation requirements does not improve outcomes. And, although the conditions in our experiment are very special (e.g. homogenous players, small groups, perfect enforcement of contributions by an institution), we believe that the basic findings

[3]The total number of institutions under the weak participation requirement is even smaller than under the unanimity rule. However, the difference is not significant.

are likely to hold in more complex settings as well.

Furthermore, in view of the debate about the importance of social preferences, it is interesting to note that a substantial proportion of groups implemented institutions of size three in treatments 2 and 3. This implies that these groups apparently tolerate that one player free rides on the public goods contributions of those who joined the institution, which in turn can be interpreted as evidence against a large proportion of inequality averse subjects.

Regarding the existing literature, our paper is related to a number of experimental papers on endogenous institution formation, most of which, however, consider the case where the institution governs all players and hence there is no second order public goods problem (Walker et al., 2000; Gürerk et al., 2006; Tyran and Feld, 2006; Kroll et al., 2007; Ertan et al., 2009; Sutter et al., 2010). Only few papers study coalition formation when the coalition only governs its members: Kosfeld et al. (2009) provide experimental evidence on the endogenous formation of a punishing institution showing that in most cases all players become members of the institution if an institution is formed at all. Dannenberg et al. (2010) compare the effectiveness of institutions that differ in the level of public good provision required from its members. The experimental data shows that weakening the rules such as to lower the free riding incentives in general does not increase overall efficiency which is akin to our result that efficiency is not increased under a weaker minimum participation rule. In a follow-up paper, Dannenberg (2011) studies coalition formation when members can vote on a binding minimum provision level of the public good and finds that social welfare is independent of the voting rule. In all of these papers players could form institutions of arbitrary size, while we set out to study institution formation under different minimum participation rules. A further interesting exception is Hamman et al. (forthcoming), who study the effects of endogenous delegation of contribution decisions to an allocator whose decision only affects those who decided to vote. They find that delegation in general improves contributions to the public good and that allowing players to communicate increases the frequency of (endogenous) delegation.

Moreover, our paper is related to Carraro et al. (2009) who analyze a model of coalition formation under a minimum participation rule which is endogenously determined by unanimity voting before the coalition formation stage.[4] The authors derive conditions on the players' payoff functions under which a particular minimum participation requirement (including the unanimity rule) is chosen in equilibrium.

The outline of our paper is as follows. In Section 2 we introduce our model of endogenous institution formation under a minimum participation rule and derive the theoretical predictions for standard and social preferences. In Section 3 we present the design of our laboratory experiment. The experimental results are provided in Section 4 and discussed in Section 5. Finally, Section 6 concludes.

2 Model and Theoretical Predictions

2.1 The Basic Model

Consider the following symmetric linear public goods game PG with $n \geq 2$ players. Each player i has a private endowment $w > 0$ and can choose to contribute $g_i \in [0, w]$ to the public good. For given contributions $g = (g_1, \ldots, g_n)$ player i's payoff is

$$\pi_i(g) = w - g_i + a \sum_{j=1}^{n} g_j \tag{1}$$

where a with $\frac{1}{n} < a < 1$ is the marginal benefit from contributions to the public good. Since $a < 1$, $g_i^0 = 0$ is a dominant strategy for all i, which implies that the game has a unique Nash equilibrium where no one contributes to the public good. However, since $na > 1$, the welfare maximizing strategy profile is $g^* = (w, \ldots, w)$.

Suppose now that before playing PG players can decide whether to join an institution that enforces full contributions to the public good.[5] For the sake of argument, we assume that the institution is costless (the subsequent results do not change as long as the welfare gain of full public good contributions outweighs the costs, though). The institution is implemented if and only if an exogenously given

[4]This modeling raises the obvious question why the unanimity voting rule is not chosen endogenously as well. It is self-evident that there is no limit to the number of stages that could be added in order to endogenize the rules under which the rules of the game are determined.

[5]We treat the institution as a black box as the specific mechanism which implements the desired contributions is of no relevance for our analysis.

minimum participation requirement is met. If the institution is implemented, all members are forced to contribute their full endowment to the public good, while all non-members can freely choose their contribution level. More precisely, for any given minimum participation requirement m with $1 \leq m \leq n$ we consider the following two-stage institution formation game IFm:

Stage 1: All players simultaneously decide whether to join the institution or not.

Stage 2: If at least m players joined the institution in stage 1, the institution is implemented. All members i of the institution are restricted to contribute their full endowment w to the public good, i.e. $g_i = w$, while all non-members j simultaneously choose their contribution $g_j \in [0, w]$. If less than m players joined the institution in stage 1, the institution is not implemented and all players i simultaneously choose their contribution $g_i \in [0, w]$. Players' payoffs are given by (1).

We also consider another institution formation game where the minimum participation requirement is weakened if no institution is implemented under a stricter minimum participation requirement. In particular, we consider the following three-stage institution formation game, IFmk, where m is the minimum participation requirement in a first participation phase and $k < m$ is the minimum participation requirement in a second participation phase:

Stage 1a: All players simultaneously decide whether to join the institution or not.

Stage 1b: If at least m players joined the institution, the game moves to stage 2. Otherwise, players again decide simultaneously whether to join the institution or not.

Stage 2: If at least m players joined the institution in stage 1a or at least k players joined the institution in stage 1b, the institution is implemented. All members i of the institution are restricted to contribute their full endowment w to the public good, i.e. $g_i = w$, while all non-members j simultaneously choose their contribution $g_j \in [0, w]$. Otherwise, if less than m players joined the institution in stage 1a and less than k players joined the institution in stage 1b, the institution is not implemented and all players i simultaneously choose their contribution $g_i \in [0, w]$. Players' payoffs are given by (1).

In the following we characterize the set of subgame perfect Nash equilibria for standard preferences as well as for social preferences as proposed by Fehr and Schmidt (1999).

2.2 Institution Formation with Standard Preferences

Under standard preferences a player's utility equals her payoff, i.e. $u_i(g) = \pi_i(g)$ for all contribution profiles $g = (g_1, \ldots, g_n)$. The following proposition provides a characterization of the pure strategy subgame perfect Nash equilibria of the institution formation games (the proof is straightforward and therefore omitted).[6]

Proposition 2.1 *Let $u_i = \pi_i$ for all players i.*

(i) Let $ma > 1$. Then in any pure strategy subgame perfect Nash equilibrium of IF*m either an institution with exactly m members is implemented or no institution is implemented.*

(ii) Let $m > k$ and $ka > 1$. Then in any pure strategy subgame perfect Nash equilibrium of IF*mk either an institution with exactly m members is implemented in stage 1a or an institution with exactly k members is implemented in stage 1b or no institution is implemented.*

Note that, by Proposition 2.1, all institution formation games have subgame perfect equilibria in which no institution is implemented. From a social planner's point of view, this might seem unsatisfactory as it implies that in theory none of the institution formation games is strictly preferable. However, the different socially undesirable equilibria are all such that all players are indifferent between joining and not joining the institution; this is due to the fact that the commitment entailed in the decision to join is only binding if sufficiently many others join as

[6] There also exist subgame perfect Nash equilibria in mixed strategies.

well. One may therefore argue that such equilibria are inherently unstable. Facing a similar problem, Kosfeld et al. (2009) focus on strict equilibria:[7]

Definition 2.1 *A subgame perfect Nash equilibrium of an extensive form game is called **stagewise strict**, if in every stage game every player's strategy is a unique best response to the equilibrium strategies of the other players.*

Using this strictness refinement the following result is straightforward.

Proposition 2.2 *Let $u_i = \pi_i$ for all players i and let $ma > 1$. Then, in any stagewise strict subgame perfect Nash equilibrium of* IFm*, an institution with exactly m members is implemented.*

Unfortunately, the strictness requirement is too strong for IFmk. In all subgame perfect Nash equilibria of IFmk no institution is formed in at least one stage of the game and hence, for at least one player the corresponding equilibrium strategy is no unique best response. Thus, there exists no stagewise strict subgame perfect Nash equilibrium in IFmk.[8]

2.3 Institution Formation with Social Preferences

Experimental research over the last years has provided ample evidence for behavior being governed by social preferences. In the context of institution formation, players with social preferences may prefer no institution over an institution that is implemented by a subgroup of players, because the latter yields unequal payoffs

[7]In the sequel, we refer to these equilibria as "stagewise strict" in order to avoid any confusion with the notion of a strict Nash equilibrium of a normal form game. Note that the notion of a stagewise strict Nash equilibrium is equivalent to the notion of a strict Nash equilibrium in the agent normal form of the extensive game.

[8]A possible way to derive a stricter prediction for $IFmk$ is to focus on cases where at least one player is not indifferent between his equilibrium strategy and some other strategy. One way to do so is to consider only subgame perfect Nash equilibria for which in every stage game there exists at least one player whose equilibrium strategy is a unique best response to the equilibrium strategies of the other players. If we apply this refinement, in the only remaining equilibria of *IFmk* an institution with exactly k members is implemented in stage 1b.

to members and non-members if non-members have no incentives to contribute to the public good. For tractability reasons and in order to compare our results to those of Kosfeld et al. (2009), we apply the social preference model by Fehr and Schmidt (1999), where players are assumed to be inequality averse.[9] The payoffs of the players are given by $\pi = (\pi_1, \ldots, \pi_n)$ as above. The utility of a player i is then defined as

$$u_i(g) = \pi_i(g) - \frac{\alpha_i}{n-1}\sum_{j\neq i}\max\{\pi_j(g)-\pi_i(g),0\} - \frac{\beta_i}{n-1}\sum_{j\neq i}\max\{\pi_i(g)-\pi_j(g),0\}. \tag{2}$$

The two parameters α_i and β_i measure the reduction of player i's utility due to disadvantageous inequality and advantageous inequality, respectively. Typically, it is assumed that $\alpha_i \geq \beta_i$ for all i and that $0 \leq \beta_i < 1$, which we also apply here.

In the sequel, we analyze the subgame perfect Nash equilibria of the institution formation games IF4, IF3 and IF43 conditional on the inequality aversion parameters of the players. To begin with, we note that, for the linear public good game with payoffs given in (1), Fehr and Schmidt (1999) show that it is a dominant strategy for each player with $a + \beta_i < 1$ to choose $g_i = 0$. Furthermore, they also prove that, if h denotes the number of players with $a + \beta_i < 1$ and the condition $\frac{h}{n-1} > \frac{a}{2}$ is met, the unique equilibrium is $g_i = 0$ for all i.

Now, suppose that there exists at least one player with $a + \beta_i < 1$ ($\iff \beta_i < 0.6$). Then, if no institution is formed, every player i gets a payoff (and utility, since there is no inequality) of $u_i = w$. If a 4-player institution is formed, the utility is $4aw = 1.6w$ for all players. By contrast, if $h = 0$, i.e. $\beta_i \geq 0.6$ for all i, there are multiple equilibria in the public good game without institution where each player contributes $g_i = g \in [0, w]$. In this case, all players i receive a payoff (and utility) of $u_i = w + g(4a - 1) = w + 0.6g$ when no institution is formed. The following result then is straightforward.

Proposition 2.3 *Let u_i be given as in (2) for all i. Then, for all types of players in any subgame perfect Nash equilibrium of* IF4*, either an institution with four members is implemented or no institution is implemented.*

[9] We note that there are other models of social preferences and more specifically inequity aversion, among them Bolton and Ockenfels (2000), Charness and Rabin (2002), and Cox et al. (2007).

As a next step, we consider the case where the minimum participation requirement is 3 (IF3 or stage 1b of IF43). In order to determine the players' utilities if a 3-player institution is implemented, we have to distinguish between the following three cases depending on the preferences of the non-member j:

- If $\beta_j < 0.6$, then j contributes $g_j = 0$ and utilities are given by $u_i = 3aw - \frac{\alpha_i}{3}w = 1.2w - \frac{\alpha_i}{3}w$ for the members and $u_j = w + 3aw - \beta_j w = 2.2w - \beta_j w$ for the non-member.

- If $\beta_j > 0.6$, then j contributes $g_j = w$ and all players receive $4aw = 1.6w$ as in the 4-player institution.

- If $\beta_j = 0.6$, the non-member j is indifferent between all contributions in $[0, w]$ and always receives a utility of $4aw = 1.6w$. For the members of the institution, this results into utilities between $3aw - \frac{\alpha_i}{3}w = 1.2w - \frac{\alpha_i}{3}w$ and $4aw = 1.6w$ depending on $g_j \in [0, w]$.

Joining a 4-player institution is only a best response for a player i with $\beta_i \geq 0.6$ since i receives a utility $w+3aw-\beta_i w = 2.2w-\beta_i w$ when deviating (i.e. not joining the institution) and $2.2w - \beta_i w > 4aw = 1.6$ if and only if $\beta_i < 0.6$. Joining a 3-player institution is a best response for player i if $3aw - \frac{\alpha_i}{3}w \geq w \Leftrightarrow \alpha_i \leq 0.6$. Thus, we can characterize the behavior of three types of subjects:

Type 1: $\beta_i \leq \alpha_i \leq 0.6$ and $\beta_i < 0.6$. This includes players with standard preferences who have parameters $\alpha_i = \beta_i = 0$. A player of type 1 is weakly averse against disadvantageous and advantageous inequality. It is a dominant strategy for this type to choose $g_i = 0$ whenever the player is not a member of an institution.

Type 2: $\alpha_i > 0.6 > \beta_i$. A player of type 2 is strongly averse against disadvantageous inequality but only weakly averse against advantageous inequality. For this type it is a dominant strategy to choose $g_i = 0$ whenever the player is not a member of an institution.

Type 3: $\alpha_i \geq \beta_i \geq 0.6$. A player of type 3 is strongly averse against disadvantageous and advantageous inequality. If $\beta_i > 0.6$, it is a dominant strategy for this type to choose $g_i = w$ whenever an institution forms and the player is not a member of the institution. If $\beta_i = 0.6$, this type is indifferent between all contributions in $[0, w]$ whenever an institution forms and the player is not a member.

Apparently, types 1 and 2 always prefer to be non-member of a 3-player institution over being a member of a 4-player institution as these types are only weakly averse against advantageous inequality. Hence, for weaker minimum participation requirements only type 3 players will form 4-player institutions. Moreover, no type 2 player would want to be part of a 3-player institution if the non-member does not contribute fully to the public good as this type is strongly averse against disadvantageous inequality. Thus, no 3-player institution will form if there is more than one player of type 2 or the non-member is not of type 3 and hence contributes fully. Type 1 players on the other hand have no reservations against being members of 3-player institutions.

Thus, using the above type-classification, we obtain the following straightforward results for IF3 and IF43.

Proposition 2.4 *Let u_i be given as in (2) for all i and consider the institution formation game* IF3*.*

(i) The implementation of a 4-player institution is supported as a subgame perfect Nash equilibrium if and only if all players are of type 3*.*

(ii) The implementation of a 3-player institution is supported as a subgame perfect Nash equilibrium if and only if either at least one player is of type 3 *or at least three players are of type* 1*.*

(iii) No institution is always supported as a subgame perfect Nash equilibrium and it is the only subgame perfect Nash equilibrium outcome if at least two players are of type 2 *and no player is of type* 3*.*

From Proposition 2.4 it follows that social preferences can be detrimental for efficiency under a weak minimum participation requirement: If no player has a strong aversion against advantageous inequality and at least two players have a strong aversion against disadvantageous inequality no institution is formed and welfare is minimized in equilibrium.

Finally, we derive the equilibrium outcomes of IF43:

Proposition 2.5 *Let u_i be given as in (2) for all i and consider the institution formation game* IF43*.*

(i) The implementation of a 4*-player institution in stage 1a is always supported as a subgame perfect Nash equilibrium.*

(ii) The implementation of a 4*-player institution in stage 1b is supported as a subgame perfect Nash equilibrium if and only if all players are of type* 3*.*

(iii) The implementation of a 3*-player institution in stage 1b is supported as a subgame perfect Nash equilibrium if and only if either at least one player is of type* 3 *or at least three players are of type* 1*.*

(iv) No institution is always supported as a subgame perfect Nash equilibrium.

2.4 Predictions

From the previous analysis, we can derive a number of testable predictions. All predictions are derived under the assumption that players either all have standard or social preferences of the Fehr and Schmidt (1999) type and that not all players are strongly averse against advantageous inequality (type 3).

Prediction 1 *The formation of an institution increases the level of public good provision.*

Concerning the average size of an institution, the theoretical analysis yields a clear prediction for IF4 and IF3 only, which we state as our second prediction:

Prediction 2 *The average size of the institutions that are implemented is larger the stricter the minimum participation requirement, i.e. the average size of an institution in* IF3 *is smaller than four.*

Concerning the number of implemented institutions and, hence, overall efficiency achieved by the different minimum participation rules, our theoretical analysis yields ambiguous predictions, which is due to the multiplicity of equilibria and the sensitivity towards the types of the players in case of social preferences. On the one hand one may predict that more institutions are implemented and hence the level of public good provision is higher under a stricter minimum participation requirement, since (1) the formation of a three-player institution involves a coordination problem, which may very likely result in miscoordination and a complete failure to implement an institution, and (2) the presence of type 2 players, who are strongly averse against disadvantageous but not against advantageous inequality, may preclude the implementation of a three-player institution. Following this line of argument we would predict that more institutions are implemented in IF4 than in IF3 and that IF43 is somewhere in between. On the other hand, under all minimum participation rules the implementation of an institution may simply fail because some players expect that everyone is playing the no institution equilibrium. This kind of miscoordination is more likely to lead to a failure to implement an institution under a strict than under a weak minimum participation requirement. If we follow this argument we would predict that more institutions are implemented in IF3 than in IF4 and that again IF43 is somewhere in between. Given these opposing effects of a change in the minimum participation rule we would predict:

Prediction 3 *The number of implemented institutions is the same for all minimum participation rules.*

From Predictions 1-3 we derive our final prediction:

Prediction 4 *The level of public good provision is larger the stricter the minimum participation requirement, i.e. there are more contributions to the public good in* IF4 *than in* IF3*.*

3 The Experiment

Our experimental design consists of the three treatments IF4, IF3 and IF43. In all treatments subjects play ten rounds of the symmetric linear public good game with endowment $w = 20$ and the marginal benefit parameter $a = 0.4$.

At the beginning of each round, subjects were asked whether they want to join an institution and thereby committing themselves to contribute their whole endowment to the public good if the institution is implemented. The three treatments differ from each other only in the minimum participation requirement of institution formation, and the subjects play the standard public goods game if no institution is implemented.

After the institution formation stage all subjects are informed whether an institution has been implemented or not. If an institution is implemented in IF4, all subjects automatically contribute their full endowment to the public good. If an institution with three members is implemented in treatments IF3 and IF43, the non-member can choose her contribution to the public good while the members of the institution only get informed that they are members of a three-player institution. At the end of each round, a summary screen is shown with the subject's individual contribution to the public good, the total group contribution and the subject's individual payoff in this round.

The experiment was conducted at the experimental laboratory of the School of Business, Economics and Social Sciences at the University of Hamburg between January and May 2011. In total, we ran nine sessions (three per treatment) with 196 mainly undergraduate students predominantly from the social sciences. In each session $20 - 24$ students participated, resulting in 17 groups for IF4 and 16 groups each for treatments IF3 and IF43. We used z-tree (Fischbacher, 2007) for programming and ORSEE (Greiner, 2004) for recruiting.

Each session started with a short introduction after which the instructions were read aloud, so that all participants knew that everyone received the same instructions. The actual experiment did not start until each subject had correctly answered a set of control questions. Subjects were then randomly matched into groups of four players who stayed together throughout the entire experiment (partner matching), however the identities of group members remained unknown.

At the end of the experiment, one of the ten rounds was chosen at random as the round determining the earnings of the subjects.[10] The exchange rate between Euro and experimental currency units (ECU) was 1 : 3. Additionally, subjects received a show-up fee of 4 Euro. The payment was carried out after a final questionnaire which included socio-demographic data and the Big-Five-Inventory-Shortversion (BFI-S). The sessions lasted $60 - 90$ minutes including instructions, control questions, questionnaire and payment. On average, a subject earned 13 Euro.

4 Results

In the sequel, we present the different results from the experiment.

4.1 Aggregate Behavior

4.1.1 Efficiency of Institutions

To begin with, we consider the general effect of institutions being present. Figure 1 depicts average contribution levels when an institution is formed and when no institution is formed. As hypothesized, for all three treatments there are more contributions to the public good when an institution is implemented. For IF4, the average contribution obviously is 20 if an institution is formed, whereas it is only approximately 5.2 when no institution is formed. IF3 records mean contributions of 16.5 if an institution is formed and 3.7 respectively if not. Subjects in IF43 on average contributed 18.6 to the public good when an institution was formed and 5.5 when no institution was formed.

These differences between contributions with and without institutions are confirmed by several Mann-Whitney tests using the average contribution per group over all periods as one independent observation. For all treatments, the differences between the contributions when an institution is implemented and when no institution is implemented are significant on a 1%-level, which confirms Prediction 1:

[10] A ten-sided dice was thrown in public by the supervisor of the experiment in order to select the payment round.

Result 1 *Average contributions to the public good are higher with than without an institution.*

4.1.2 Comparing Different Minimum Participation Requirements

Figure 1 also shows that average contributions to the public good when no institution is formed are only slightly different in the three treatments (5.2, 3.7 and 5.5). This observation is confirmed by Mann-Whitney tests, which report no significant differences. However, Figure 2 paints a different picture when all contributions are taken into account. In total, the average contributions are 13.6 in IF4, 12.4 in IF43 and 9.3 in IF3 respectively, indicating that IF4 and IF43 produce higher contributions to the public good than IF3 (Mann-Whitney test 5%-level). Yet, although the provision level is slightly higher in IF4 than in IF43, we do not find significant differences in the public good provision between these two treatments ($p = 0.43$).

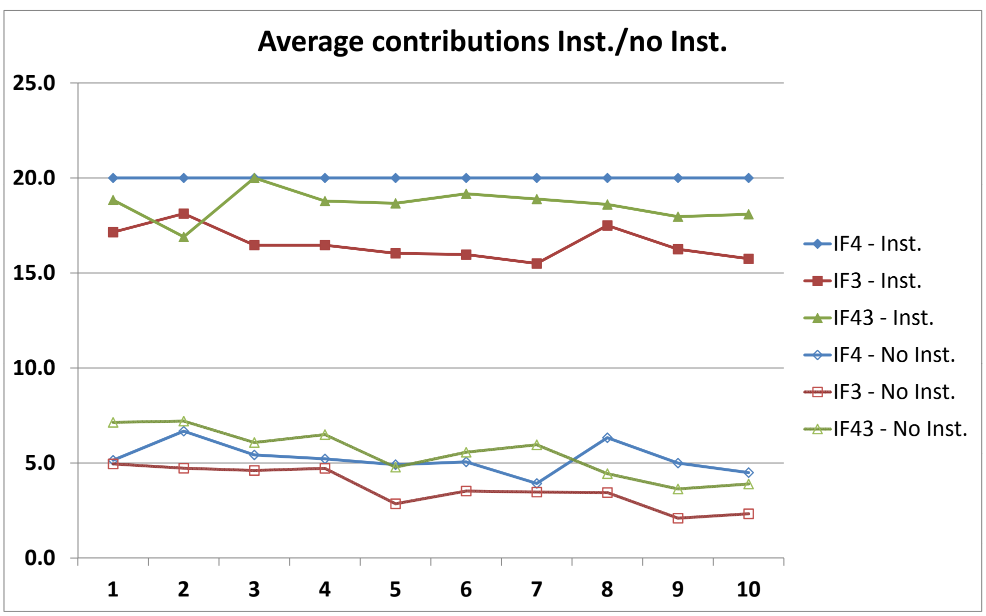

Figure 1: Efficiency of institution formation in periods 1-10.

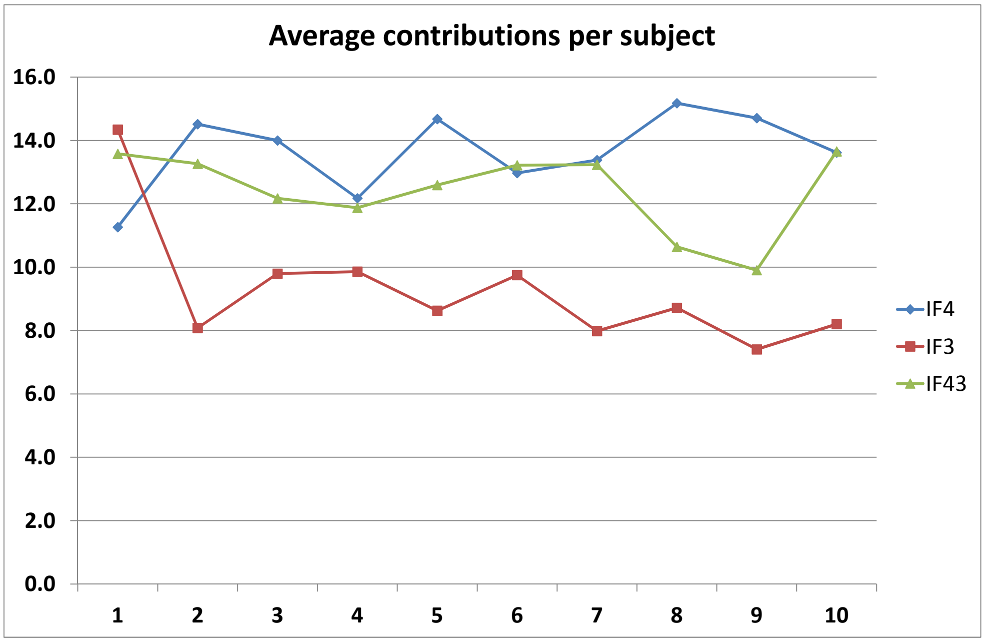

Figure 2: Average contributions across treatments in periods 1-10.

Result 2 *Contributions to the public good are higher in* IF4 *and* IF43 *than in* IF3*, while there is no significant difference between* IF4 *and* IF43*. When no institution is implemented, all treatments lead to the same level of public good provision.*

These results are confirmed by the regression depicted in Table 1, where IF4 is used as the baseline treatment. We also included age, sex and the big five personality traits from the questionnaire in the regression. However, the effects on the coefficients for IF3 and IF43 are negligible and in all our regressions the socio-demographic variables and BIG5 personality traits show no consistent and mostly insignificant effects on the behavior of the subjects. We therefore do not report the coefficients of these variables in our tables. Table 1 shows that, in accordance with our prediction, the contributions to the public good are significantly lower in IF3 than in IF4 (1% significance). Contributions are also lower in IF43 than in IF4 but the coefficient is not significantly different from zero. These observations confirm Prediction 4 and in addition show that overall efficiency is the same in IF4 and IF43, which we could not derive from our theoretical analysis.

Ind. Variable	PG Contribution
Const.	9.416^{**}
	(4.157)
IF3	-4.382^{***}
	(1.450)
IF43	-1.132
	(1.643)
Observations	$N = 1960$

Table 1: GLS regression results for public goods contributions across treatments. Contributions in IF4 are baseline. IF3 (IF43) is a dummy variable that takes the value 1 if the treatment is IF3 (IF43) and 0 otherwise. Std. errors in parentheses are adjusted for group clusters. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

A straightforward conjecture is that Result 2 can be explained by the number and size of the institutions implemented in the different treatments. Table 2 reports the percentages of different institution sizes across treatments. We find that the observed differences in contributions to the public good cannot be explained with the total number of formed institutions. Although there are more institutions implemented in IF4 (institutions implemented in 57% of the cases) than in IF43 (53%) and IF3 (44%), these differences are not significant.[11] Thus, the following result confirms Prediction 3:

Result 3 *There is no difference in the number of institutions implemented in all treatments.*

[11]For these tests, the number of institutions per group were used as one independent observation.

Inst. Size	**IF**4	**IF**3	**IF**43
4	57%	9%	35%
3	0%	35%	18%
Total	57%	44%	53%

Table 2: Size of institutions (proportion of institutions of size 4 and 3, respectively, taken over all groups and all periods).

We continue by comparing the different sizes of institutions across treatments. The percentages of institutions with four members are 35% in IF43 and 9% in IF3. Again applying Mann-Whitney tests, we can establish that there are more institutions with four members in IF4 than in IF43 (10% significance) and IF3 (1% significance). Also, groups form 4-player institutions more frequently in IF43 than in IF3 (5% significance). Yet, there are more institutions with three members in IF3, 35%, than in IF43, 18% (1% significance). Overall, Mann-Whitney tests show that the average size of an institution formed in IF4 is larger than the average size of an institution in IF3 or IF43 (1% significance each). Furthermore, the average size of an institution in IF43 is larger than in IF3 (5% significance). This confirms Prediction 2 and even extends it to IF43. We summarize in our next result:

Result 4 *On average, the largest institutions are formed in* IF4, *and institutions are larger in* IF43 *than in* IF3.

Moreover, the contribution of a non-member when an institution of size 3 has been formed does not differ between IF3 and IF43, as the non-member on average contributes 2.5 to the public good in both treatments. A Mann-Whitney test confirms that in IF3 the contribution of a non-member when an institution is implemented is significantly lower than the contribution if no institution has been formed (5% significance). In IF43, however, there is no such difference.

We explore the institution formation in IF43 in more detail. Figure 3 depicts the proportion of cases in which an institution of size three or four is formed in stage 1a and 1b, respectively. 62% of the institutions are formed already in stage 1a and only 38% in stage 1b. A Mann-Whitney test confirms that this difference is significant (5%-level). Furthermore, three institutions of size 4 were even formed in stage 1b, which implies that overall 66% of all implemented institutions are of size 4.

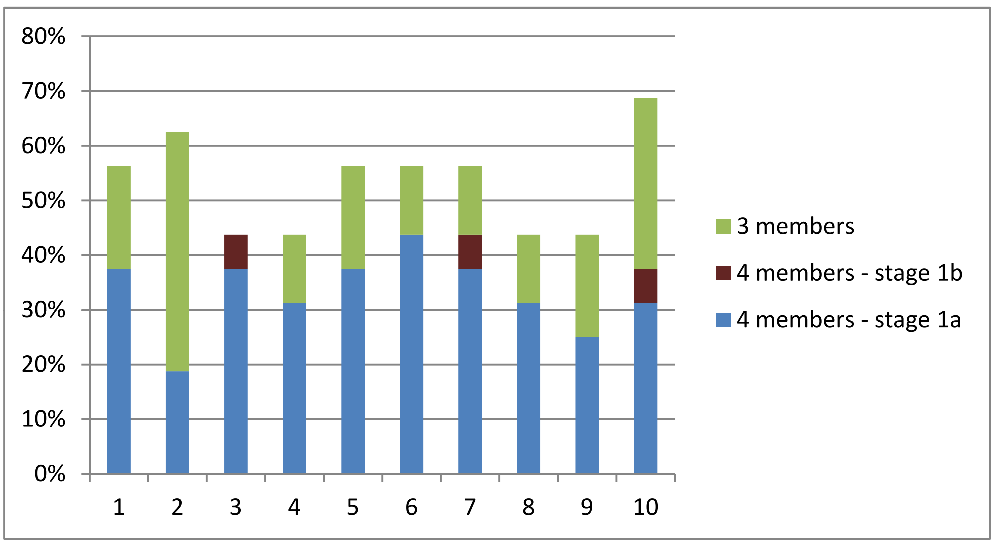

Figure 3: Institutions formed in IF43 in periods 1-10 (proportion of cases with an institution of size 3 or 4 in stage 1a and 1b).

Summarizing, the observations concerning the number and size of institutions support the findings regarding the contributions across treatments: While there is no difference in the total number of institutions in IF3 and IF43, the latter has significantly more 4-player institutions leading to higher average contributions in IF43 than in IF3. There are also more 4-player institutions in IF4 than in IF43 but the difference is only weakly significant. This, together with the fact that there is also a number of 3-player institutions in IF43 may explain why we do not observe a significant difference in the contributions between IF4 and IF43.

4.2 Individual Behavior

Having analyzed the differences in aggregate outcomes across treatments, we now proceed to explore individual behavior under the different minimum participation requirements.

4.2.1 Decision to Join the Institution

In order to further understand the participation decisions of the subjects across treatments, we performed two probit regressions. The results are presented in Table 3. The regressions depict the impact of the different treatments on the probability to join the institution. The two regressions differ in the way that the "join" decision is defined for IF43; in (1) only stage $1a$ is considered, in (2) the ultimate decision in each period is used. That is, if no institution has been formed in stage $1a$, then the decision in stage $1b$ is used in the regression. We find that in the treatments IF3 and IF43 the probability that a subject joins the institution is lower compared to IF4 (1% significance), and that the probability is even less for IF3 than for IF43 (5% significance).[12] Regression (2) shows that the willingness to join the institution in IF43 is less when also considering the decision in stage $1b$; it is then not significantly higher than in IF3.[13]

Result 5 *The probability that a subject joins an institution is higher in* IF4 *than in* IF3 *and* IF43.

Observe that Result 5 is consistent with Results 3 and 4: The higher the probability to join an institution the higher the probability that a 4-player institution is formed. On the other hand, for given probabilities to join an institution, the probability that an institution is formed decreases with the minimum participation requirement. In our case the two effects of a higher probability to join an institution and a lower probability that an institution is implemented at all cancel out, so that there is no difference in the number of institutions formed in the different treatments.

[12]This regression with IF3 as baseline treatment is not shown in a table. The coefficient is 0.634 for IF43 with standard deviation 0.255 and $p = 0.013$.

[13]This regression is also not shown in a table. The coefficient for IF43 is 0.264 and $p = 0.256$ with again IF3 as baseline treatment.

	Join Institution	
Ind. Variable	(1)	(2)
Const.	1.852^{***}	1.796^{***}
	(0.214)	(0.202)
IF3	-1.476^{***}	-1.428^{***}
	(0.280)	(0.264)
IF43	-0.833^{***}	-1.163^{***}
	(0.279)	(0.263)
Observations	$N = 1960$	$N = 1960$

Table 3: Probit regression results for the probability to join the institution across treatments. Regressions (1) and (2) differ in the definition of the participation decision in IF43. In regression (1) we use the choice in stage $1a$, whereas in (2) we apply the choice in the ultimate stage, i.e., the decision of stage $1a$ if an institution was implemented in stage $1a$ and the decision of stage $1b$ otherwise. The decision to join the institution in IF4 is baseline. IF3 (IF43) is a dummy variable that takes the value 1 if the treatment is IF3 (IF43) and 0 otherwise. Std. errors in parentheses adjusted for group clusters. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

We again used a Mann-Whitney test to compare IF3 with stage 1b of IF43. For this test, the number of institutions per group is not suitable to count as an observation, as the groups in IF43 do not always reach the second stage. Therefore, we used the number of institutions per group per stage 1b played as one independent observation.[14] We find that there are more institutions formed in IF3 than in stage 1b of IF43 (Mann-Whitney test 5%-level). Hence, the willingness to join the institution is higher in IF3 than in stage 1b of IF43. One interpretation of this finding is that the subjects in IF43 interpret the failure to form an institution in stage 1a as a signal that at least one player will free ride on the public good provision of the others which may induce subjects with social preferences not to join the institution in the second stage.

To test this and to get an understanding of the subjects' individual participation decisions, we performed separate probit regressions for all treatments. These regression results are summarized in Table 4. As treatment IF43 has two participation stages, we performed regressions for each of the possible stages 1a and 1b. We find that for all treatments the decision to join the institution is largely influenced by the decision made in the previous period, i.e. a subject is more likely to join the institution if it had joined already in the previous period. For IF43 this property needs to be specified for the different stages, though. In stage 1a, the decision of stage 1a of the previous period has an effect, while the decision in stage 1b of the previous period does not. Similarly, for the decision in stage 1b, only the decision in stage 1b of the previous period has a significant influence on the choice. In stage 1b of IF43, the decision to join the institution in stage 1a increases the probability to join the institution also in stage 1b.

Furthermore, we observe that in IF3 and in stage 1b of IF43 a 3-player institution in the previous period has a negative impact on the subjects' decision to join the institution. We included a dummy variable for the membership in a 3-player institution in the previous period in the regressions for IF3 and IF43. However, this variable and the dummy variable for a 3-player institution in the previous period are highly correlated ($|\rho| \geq 0.79$ in each treatment) and we therefore only report Inst3_{t-1} in the regressions.

[14] The played stages vary from one up to ten across the different groups.

	Join Institution			
Ind. Variable	IF4	IF3	IF43-1a	IF43-1b
Const.	-0.607	-4.587^{***}	-0.375	-2.305^{*}
	(1.306)	(1.514)	(1.391)	(1.319)
Join1_t				1.772^{***}
				(0.225)
Join1_{t-1}	0.853^{***}	1.210^{***}	0.594^{***}	0.231
	(0.282)	(0.220)	(0.228)	(0.253)
Join2_{t-1}			0.013	1.140^{***}
			(0.226)	(0.301)
Inst4_{t-1}	0.406^{*}	-0.669^{**}	-0.005	-0.025
	(0.233)	(0.284)	(0.230)	(0.255)
Inst3_{t-1}		-0.935^{***}	-0.290	-1.332^{***}
		(0.302)	(0.183)	(0.215)
Observations	$N = 612$	$N = 576$	$N = 576$	$N = 388$

Table 4: Probit regression results for the probability to join the institution in the different treatments. The decision not to join the institution is baseline. Join1_t is a dummy variable for IF43 that takes the value 1 if the subject joined the institution in stage 1a of the current period and is 0 otherwise. Join1_{t-1} is a dummy variable that takes the value 1 if the subject joined the institution in the previous period (in stage 1a of the previous period for IF43) and is 0 otherwise. Join2_{t-1} is a dummy variable for IF43 that takes the value 1 if the subject joined the institution in stage 1b of the previous period and is 0 otherwise. Inst4_{t-1} is a dummy variable that takes the value 1 if a 4-player institution was implemented in the previous period and is 0 otherwise. Inst3_{t-1} is a dummy variable that takes the value 1 if a 3-player institution was implemented in the previous period and is 0 otherwise. Std. errors in parentheses adjusted for group clusters. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

4.2.2 Learning

In order to determine if there is any learning over the ten periods, we used the Spearman-Rank-Order test. We find that there is no correlation between contributions to the public good and the periods in IF4 (Spearman-$\rho = 0.4$) and IF43 ($\rho = -0.2$). Yet, we find that there is a small negative correlation (10% significance, $\rho = -0.6$) for IF3. However, this significance vanishes if we exclude period 1. For the number of institutions formed per period, there is a slightly significant positive correlation for IF4 ($\rho = 0.55$, 10% significance). The other two treatments show no correlations ($\rho = -0.27$ for IF3 and $\rho = -0.01$ for IF43). Furthermore, there is no significant correlation between the number of choices to join an institution and the period. All results are quite robust against start-game and end-game effects, i.e. we obtain the same results if we only consider periods 3 through 8 or 4 through 7.

Result 6 *In no treatment there is a significant change in behavior over time.*

4.2.3 Incentives to Opt for Weaker Minimum Participation Rules

Finally, we were interested in the empirical incentives for subjects to free ride. More specifically, while Result 2 shows that on average players' payoffs are highest in IF4 and IF43, i.e. whenever the institution formation game starts with a strict minimum participation requirement, one may ask whether nevertheless certain player behavior generates higher payoffs under a weaker minimum participation requirement. In particular, it could be that a player who intends to free ride on the institution formation of others earns a higher payoff in IF3 and in IF43 than in IF4. However, it turns out that this is not the case. The average payoff of a player in IF4 is 28.19. Not surprisingly, the payoff conditional on joining the institution in IF4 is even higher, namely 29.00. By comparison, the average payoff of a player who does not join the institution (a free rider) is 27.09 in IF3 and 27.23 in IF43. Although the differences in payoffs are not significant, this shows that free-riders do not profit from the weakening of the minimum participation requirement in IF3 or IF43.

Result 7 *There is no significant difference between the average payoff of a free rider in* IF3 *and* IF43 *and a player in* IF4.

5 Discussion

In the following, we contrast the results derived in Section 2.3 with the data from our experiment. To begin with, we consider treatment IF3.

The first thing to note is that in IF3 we do not observe any significant contributions by non-members in case of a 3-player institution, suggesting that there are only few subjects with a strong aversion against advantageous inequality (type 3 players). Moreover, there is only one group which formed 4-player institutions in six out of ten periods as shown in Figure 4, which is only an equilibrium if all players are of type 3. In all other groups, 4-player institutions were broken up immediately in the following period, suggesting that these 4-player institutions were rather a result of a coordination problem. At the same time, we observe a substantial number of 3-player institutions (cf. Table 2) from which we conclude that less than half of all subjects are of type 2 as, by Proposition 2.4, that would prevent 3-player institutions from being formed if there are no type 3 players.

Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	3	0	4	3	3	0	3	4	3	4	3	0	4	0	3	3
2	0	0	3	0	3	0	0	0	0	3	0	0	4	0	0	0
3	3	0	0	3	3	0	4	3	0	3	0	0	4	0	0	0
4	0	0	3	0	0	4	0	3	3	0	0	0	4	0	3	3
5	0	3	0	0	0	3	3	0	0	3	3	0	4	0	3	0
6	3	0	3	0	3	0	0	3	0	3	3	0	3	0	0	3
7	3	3	0	0	3	0	0	0	0	3	0	0	3	0	0	3
8	0	4	3	0	0	0	4	3	3	0	0	4	0	0	0	0
9	0	0	3	3	0	0	3	0	0	0	0	3	4	0	3	0
10	3	3	3	0	3	0	3	3	0	0	3	0	0	0	0	0

Figure 4: Institutions formed by the groups in IF3 (0 = no institution, 3 = 3-player institution, 4 = 4-player institution).

More specifically, there was one group which did not form any institution at all in ten periods. We could therefore conclude that there was no subject of type 3 and at least two subjects of type 2 in this group, as "no institution" then is the only subgame perfect Nash equilibrium. And in fact a detailed analysis of the individual decisions reveals that two subjects never decided to join the institution. Not joining becomes the dominant strategy for a type 2 player if he or she believes that there is no type 3 player in the group. Accordingly, we detected 10 more subjects who joined the institution in at most two periods. Thus, in total we can identify 12 subjects (19%) whose behavior is consistent with a type 2 player.

Classifying the remaining 48 subjects who participated in treatment IF3 proves to be more difficult. However, there is evidence that the majority of these subjects are of type 1, as, by Proposition 2.4, in the absence of type 3 players 3-player institutions are only equilibrium outcomes when at least three players are of type 1. Thus, in the 12 groups which repeatedly form 3-player institutions, we may classify at least 36 subjects (56%) as type 1. The type of the remaining 12 subjects (19%) remains unknown from the data.

We performed a similar analysis for IF43. Just as in IF3, there is one group that forms 4-player institutions in stage 1*b* (cf. Figure 5). However, other than that there is no direct evidence of a type 3 player. Thus, we again conclude that this type is rather scarce. Moreover, we observe multiple 3-player institutions in eight groups, suggesting at least three players of type 1 in each group. Besides these 24 type 1 and four type 3 subjects, we were able to identify 16 subjects potentially of type 2 by their behavior in stage 1*b*. All of these 16 subjects chose not to join the institution in stage 1*b* in at least 80% percent of the cases.[15]

Regarding the types of the remaining 20 subjects, we can only speculate as both a 4-player institution and "no institution" are equilibrium outcomes for all group compositions. Yet, as repeatedly mentioned above, we find only little evidence for type 3 players. By contrast, we believe that type 2 subjects, having correct expectations about the types (and behavior) of the other group members, should be identifiable in stage 1*b* by their best response not to join the institution. Therefore, we believe that the 8 subjects who repeatedly reach stage 1*b* do not have a strong aversion against disadvantageous inequality and, thus, can be classified as type 1. This leaves us with a similar type distribution as in IF3. Table 5 summarizes the analysis for both treatments.

[15]For this analysis, we excluded those three groups which only reach the stage once or twice.

Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4	4	4	0	0	3	0	3	3	0	4	4	0	0	4	0
2	4	3	0	3	0	3	0	0	3	3	0	4	0	3	4	3
3	0	4	4	0	4	0	0	4	0	0	4	4	0	0	4	0
4	4	0	4	0	0	3	0	0	3	0	4	4	0	0	4	0
5	0	4	4	0	0	0	0	4	0	4	4	3	0	0	4	3
6	4	0	4	0	4	4	0	0	0	3	4	3	0	4	4	0
7	0	3	4	3	4	4	0	4	0	0	4	0	0	4	4	0
8	0	4	4	3	3	0	0	0	4	0	4	0	0	0	4	0
9	0	0	4	0	0	0	3	0	3	0	4	4	3	0	4	0
10	4	0	4	3	4	4	0	4	3	3	0	3	3	4	0	0

Figure 5: Institutions formed by the groups in IF43 (0 = no institution, 3 = 3-player institution, 4 = 4-player institution). Dark background marks 4-player institutions formed in stage 1*b*.

	IF3	**IF**43
type 1	36 (56%)	32 (50%)
type 2	12 (19%)	16 (25%)
type 3	4 (6%)	4 (6%)
unknown	12 (19%)	12 (19%)

Table 5: Estimation of the type distribution in IF3 and IF43. The total number of subjects is 64 in both treatments, percentages in parentheses.

6 Conclusion

In this paper, we have studied the effectiveness of different minimum participation rules for the formation of an institution to provide a public good. Due to a multiplicity of equilibria it is not clear which rule is optimal from a theoretical point of view: All minimum participation rules allow for a subgame perfect Nash equilibrium, where the formation of an institution fails and hence welfare is minimized. This is true independent of whether players have standard preferences or Fehr and Schmidt (1999) preferences, i.e. they are inequality averse.

The intuitive conjecture, that in practice a weaker minimum participation requirement increases the number of institutions and hence increases efficiency is not confirmed by our laboratory experiment. In fact, the total number of institutions is independent of the minimum participation rule. While more large institutions are formed under the strict rule that requires the participation of all players, small (3-player) institutions are frequently formed if the minimum participation requirement is weakened either from the very beginning or else in a second stage. Overall efficiency in the experiment is highest in those treatments that start with or are restricted to the strict unanimity rule.

Moreover, as we have seen, observed behavior in the experiment is consistent with a majority of subjects having standard preferences and at most one third being strongly inequality averse. In that sense, our results stand in contrast to the findings by Kosfeld et al. (2009) who find that most players are not willing to implement institutions that do not involve all group members which can only be explained if most players are strongly inequity averse.

While we are hesitant to draw any bold policy conclusions from our small scale laboratory experiment, we nevertheless believe that our results support the idea that using strict minimum participation rules is best whenever it comes to the implementation of a policy, e.g. a follow-up agreement to the Kyoto protocol. In particular, the fear that a strict rule exacerbates the implementation of an institution appears to be unwarranted.

References

Abreu, D., and A. Sen (1990): "Subgame Perfect Implementation: A Necessary and Almost Sufficient Condition," *Journal of Economic Theory*, 50, 285–299.

Barrett, S. (1994): "Self-Enforcing International Agreements," *Oxford Economic Papers*, 46, 878–894.

Bolton, G., and A. Ockenfels (2000): "ERC - A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, 90, 166–193.

Carraro, C., and C. Marchiori (2003): "Stable Coalitions," in *The Endogenous Formation of Economic Coalitions*, ed. by C. Carraro, chap. 5, pp. 156–198. Edward Elgar, Cheltenham, UK.

Carraro, C., C. Marchiori, and S. Oreffice (2009): "Endogenous Minimum Participation in International Environmental Treaties," *Environmental Resource Economics*, 42, 411–425.

Carraro, C., and D. Siniscalco (1993): "Strategies for the International Protection of the Environment," *Journal of Public Economics*, 52, 309–328.

Charness, G., and M. Rabin (2002): "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117, 817–869.

Cox, J. C., D. Friedman, and S. Gjerstad (2007): "A Tractable Model of Reciprocity and Fairness," *Games and Economic Behavior*, 59, 17–45.

Dannenberg, A. (2011): "Coalition Formation and Voting In Public Goods Games," *Strategic Behavior and the Environment*, forthcoming.

Dannenberg, A., A. Lange, and B. Sturm (2010): "On the Formation of Coalitions to Provide Public Goods - Experimental Evidence from the Lab," Working Paper No. 15967, NBER.

d'Aspremont, C., J. J. Gabszeweiz, and J. A. Weymark (1993): "On the Stability of Collusive Price Leadership," *Canadian Journal of Economics*, 16, 17–25.

Ertan, A., T. Page, and L. Putterman (2009): "Who to punish? Individual Decisions and Majority Rule in Mitigating the Free Rider Problem," *European Economic Review*, 53, 495–511.

Falkinger, J. (1996): "Efficient Private Provision of Public Goods by Rewarding Deviations from Average," *Journal of Public Economics*, 62, 413–422.

Fehr, E., and S. Gächter (2000): "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90, 980–994.

——— (2002): "Altruistic Punishment in Humans," *Nature*, 415, 137–140.

Fehr, E., and K. M. Schmidt (1999): "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114(3), 817–868.

Finus, M., and B. Rundshagen (2003): "Endogenous Coalition Formation in Global Pollution Control: A Partition Function Approach," in *The Endogenous Formation of Economic Coalitions*, ed. by C. Carraro, chap. 6, pp. 199–243. Edward Elgar, Cheltenham, UK.

Fischbacher, U. (2007): "Z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10, 171–178.

Gerber, A., and P. C. Wichardt (2009a): "Providing Public Goods in the Absence of Strong Institutions," *Journal of Public Economics*, 93, 429–439.

——— (2009b): "On the Private Provision of Intertemporal Public Goods with Stock Effects," *available at SSRN: http://ssrn.com/abstract=1655747*.

Greiner, B. (2004): "An Online Recruitment System for Economic Experiments," in *Forschung und wissenschaftliches Rechnen 2003*, ed. by K. Kremer, and V. Macho, vol. 63, pp. 79–93. GWDG Bericht 63, Göttingen: Ges. für Wiss. Datenverarbeitung.

Groves, T., and J. Ledyard (1977): "Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Problem," *Econometrica*, 45, 783–809.

Gürek, O., B. Irlenbusch, and B. Rockenbach (2006): "The Competitive Advantage of Sanctioning Institutions," *Science*, 312, 108–111.

Hamman, J., R. Weber, and J. Woon (forthcoming): "An Experimental Investigation of Electoral Delegation and the Provision of Public Goods," *American Journal of Political Science*.

Jackson, M. (1992): "Implementation in Undominated Strategies: A Look at Bounded Mechanisms," *Review of Economic Studies*, 59, 757–775.

Kosfeld, M., A. Okada, and A. Riedl (2009): "Institution Formation in Public Goods Games," *American Economic Review*, 99, 1335–1355.

Kroll, S., T. L. Cherry, and J. F. Shogren (2007): "Voting, Punishment and Public Goods," *Economic Inquiry*, 45, 557–570.

Masclet, D., C. Noussair, S. Tucker, and M.-C. Villeval (2003): "Monetary and Non-Monetary Punishment in the Voluntary Contributions Mechanism," *American Economic Review*, 93, 366–380.

Moore, J., and R. Repullo (1988): "Subgame Perfect Implementation," *Econometrica*, 56, 1191–1220.

Palfrey, T., and S. Srivastava (1991): "Nash Implementation Using Undominated Strategies," *Econometrica*, 59, 479–502.

Rutz, S. (2001): "Minimum Participation Rules and the Effectiveness of Multilateral Environmental Agreements," Working Paper No. 01/22, Centre for Economic Research, ETHZ, Zurich.

Samuelson, P. A. (1954): "The Pure Theory of Public Expenditure," *The Review of Economics and Statistics*, 36, 387–389.

Sefton, M., R. Shupp, and J. M. Walker (2007): "The Effect of Rewards and Sanctions in Provision of Public Goods," *Economic Inquiry*, 45, 671–690.

Sutter, M., S. Haigner, and M. G. Kocher (2010): "Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations," *Review of Economic Studies*, 77, 1540–1566.

Tyran, J. R., and L. P. Feld (2006): "Achieving Compliance when Legal Sanctions are Non-deterrent," *Scandinavian Journal of Economics*, 108, 135–156.

Walker, J. M., R. Gardner, A. Herr, and E. Ostrom (2000): "Collective Choice in the Commons: Experimental Results on Proposed Allocation Rules and Votes," *Economic Journal*, 110, 212–234.