Essays on the Conditional Contribution Mechanism for Public Good Provision

Dissertation

zur Erlangung des akademischen Grades
Doctor Rerum Politicarum

An der
Fakultät für Wirtschafts- und Sozialwissenschaften
der Ruprecht-Karls-Universität Heidelberg

vorgelegt von

Andreas Reischmann

Heidelberg, März 2015

To Chloe

Acknowledgments

First and foremost, I want to thank my supervisor Jörg Oechssler. Jörg, you managed to create a work environment based on trust and freedom that made my research possible. Your advise and your comments were invaluable for my work. Thank you further for granting me the opportunity to present my papers on international conferences. My doctorate under your supervision is an experience that I will value for the rest of my live. Second, I thank the secretary of our department Gabi Rauscher. Gabi, whenever I was in need of something, when I had any questions, or when I had to hand in bureaucratic papers, you were always there to help and support me. Your help made my work a lot easier. Third, I want to thank all members of the department of economics in Heidelberg as a collective. I enjoyed the extraordinary spirit, the lunch discussions, as well as our weekly badminton matches immensely. I further thank many members of the department individually for helpful discussions of my work, the time spent proofreading my papers, or advise concerning my laboratory experiments. Here I want to mention in particular Christoph Vanberg, Christoph Brunner, Hannes Koppel, Peter Dürsch, Jonas Hedlund, Holger van Schoor, and Florian Kauffeld. Next I thank the members of the CDSE in Mannheim for granting me the opportunity to attend their Ph.D. courses. The knowledge they taught me set the ground that my work is build on. Many thanks also to Alvin Roth. At the very beginning of my doctorate, during your visit in Heidelberg, you took your time to discuss my research ideas with me. The experience that the current Noble price winner in economics was ready to discuss my ideas gave me a boost of confidence that lasts until now. Finally, I want to thank my family. First, I thank my parents for helping me become the person that I am. Your education and support make my success possible. And most of all I thank my own family, my wife Chloe and our daughter Ella. Chloe, thank you for your unconditional support, for your advise, for your love. I would not be who I am without you. Ella, thank you for being. You are just days old, but your were the greatest motivation anyone can have when writing their thesis.

Contents

Chapter 1

The Conditional Contribution Mechanism for the Provision of Public Goods.*

Abstract

Many mechanisms have been designed to solve the free-rider problem in public good environments. The designers of those mechanisms focused on good static equilibrium properties. In this paper, I propose a new mechanism for the provision of public goods that has good dynamic properties instead. The mechanism gives all agents the possibility to condition their contribution on the total level of contribution provided by all agents. Under a reasonable variant of Better Response Dynamics all equilibrium outcomes are Pareto efficient. This makes the mechanism particularly suited for repeated public good environments. In contrast to many previously suggested mechanisms, it does further not require an institution that has the power to enforce participation and/or transfer payments. Neither does it use any knowledge of agents preferences.

*I thank the participants of the Game Theory Festival 2013 at Stony Brook University, as well as the participants of the IWGTS at the University of Sao Paulo. I also thank Alvin Roth, Jörg Oechssler, Christoph Vanberg, Christoph Brunner, Jonas Hedlund, Thomas Tröger and Hannes Koppel as well as all participants of the seminars at Heidelberg University for helpful comments.

1.1 Introduction

Numerous mechanisms have been developed in an attempt to solve the free-rider problem in public good scenarios. All those mechanisms were developed with a static solution concept in mind. However, Healy (2006) shows that in repeated public good environments agents' actions can be well described by a dynamic better response behavior. This paper therefore presents a new mechanism that achieves efficient contribution levels under an adjusted better response dynamic. This mechanism is called the Conditional Contribution Mechanism (CCM).

In the CCM agents can free-ride and contribute unconditionally as in the Voluntary Contribution Mechanism. Moreover, agents have the possibility to conditionally contribute. In the most simple environment contribution is binary and agents' utility from the public good increases linearly with the level of the public good. In this environment an offer of conditional contribution has the form "I am willing to contribute, if at least k agents contribute in total". The mechanism then chooses the highest possible level of total contribution that satisfies all those conditions.

Under Better Response Dynamics agents switch only to messages with positive probability that make them weakly better off if nobody else switches as well. In the proposed mechanism all agents are indifferent between many of their messages. Thus, Better Response Dynamics are not sufficiently restrictive for the dynamic process to converge to any equilibria.

However, the conditional contribution structure of the mechanism makes some better responses more plausible than others in the long term. Once a certain level of contributions is reached, messages can be separated in two sets. Under the messages in one set other agents can change the outcome such that the first agent is worse off, but still has to contribute to the public good. The second set of messages makes sure that in all outcomes in which the first agent has to contribute he is at least as well off as he is in the current outcome.

The first kind of messages, which increase other agents incentives to free-ride, shall be called exploitable. The second kind shall be called unexploitable. The solution concept used in this paper, Unexploitable Better Response Dynamics, assumes that in the long run agents only choose strategies which are better responses and not exploitable.

The central result of the binary model is that an outcome is an equilibrium outcome of the proposed mechanism under Unexploitable Better Response Dynamics if and only if it is Pareto optimal and a strict Pareto improvement over the outcome with zero contribution.

The remaining parts of the paper generalize the environment. First, contributions can now be non-binary. Here the mechanism needs to be adjusted. However, the general idea of offering agents the options to free-ride, conditionally contribute, and unconditionally contribute remains unchanged. In this environment the equilibrium results mirror the results for binary contribution.

Second, the environment is generalized to cover weakly monotonic increasing instead of linear valuation functions. In this case Pareto optimality will not be enough to ensure that an outcome is part of a recurrent class. Since utility gained from the public good no longer increases linearly with the contribution towards the public good, there might now be coalitions of agents who benefit from reducing their own contributions even if all other agents then contribute nothing any more. In this environment an outcome is an outcome of a recurrent class of the mechanism under Unexploitable Better Response Dynamics if and only if it is in the core and any deviation of a coalition from this outcome makes at least one agent in that coalition strictly worse off. This holds if at least one such outcome exists. Existence can be guaranteed by adding only infinitesimal monetary incentives.

1.1.1 Related literature

This work relates in particular to three branches of the literature. The first one is given by work on mechanisms to increase contributions to public goods. The earliest work dates back to Lindahl (1919). However, his pricing system turned out to be not incentive compatible. The most prominent incentive compatible mechanisms were then designed by Clarke (1971) and Groves and Ledyard (1977). More recent advances are the Jackson-Moulin mechanism (Jackson and Moulin, 1992) or the Falkinger mechanism (Falkinger et al., 2000). However, all those mechanisms have their own draw-backs. Some e.g. require participation to be enforceable, or a high level of information about other agents' preferences to reach the desired equilibrium.

Second, there are experimental studies on public good provision. For a general survey I refer to Ledyard (1994), or the more recent surveys of Chen (2008) and Chaudhuri (2011). Two smaller branches of this literature support the idea that the CCM should be successful.

First, the studies of Fischbacher et al. (2001) and Kocher et al. (2008) show that some agents have strict preferences for conditional cooperation. I do not use this fact in the equilibrium analysis. However, it is obvious that preferences for conditional cooperation make it more likely that agents choose to conditional contribution instead of free-riding.

Second, there are certain papers that compare the performance of the Voluntary Contribution Mechanism (VCM) experimentally to the performance of other simple public good mechanisms. Two mechanisms have been found to be able to increase contributions at least in some situations. The auction mechanism by Smith (1979, 1980) and the Provision Point Mechanism (PPM) studied e.g. in Rondeau et al. (1999, 2005). Those mechanisms have in common that they use a sharp discontinuity to prevent the incentives of free-riding. The CCM shares this property of a sharp discontinuity.

In the PPM and the auction mechanism the value of this discontinuity has to be chosen by the mechanism designer. With a lack of knowledge of agents' preferences this can lead to failure of the mechanism to provide any contributions. In the CCM the value of the discontinuity depends on agents' messages. Thus, it is no longer exogenously fixed but can dynamically adjust itself to the optimal value.

The third branch of the literature focuses on Better Response Dynamics in mechanisms. I already mentioned that Healy (2006) provides experimental evidence that agents' behavior in public good mechanisms can be well described by a better response model. The importance of Better Response Dynamics in mechanisms is further highlighted by the recent introduction of Better Response Dynamics into the implementation literature by Cabrales and Serrano (2011).

1.1.2 Plan of the paper

The remaining sections are structured as follows. In section 1.2, I introduce the Binary Conditional Contribution Mechanism (BCCM) in the simplest possible setting. Valuations are linear and contribution to the public good is binary. Section 1.3 introduces

Unexploitable Better Response Dynamics and the outcomes of recurrent classes of the BCCM under UBRD are calculated. Section 1.4 removes the assumption that contributions are binary and introduces the (non-binary) Conditional Contribution Mechanism. In Section 1.5, the assumption of linear valuations is replaced with the weaker assumption of weakly increasing valuation functions. Section 1.6 provides a summary and discussion of the results. Proofs of all theorems can be found in Appendix A.

1.2 The Binary Conditional Contribution Mechanism

I consider a public good environment in the following form. All $n \in \mathbb{N}$ agents labeled i are considered to have one monetary unit available in each period, which they can either keep or invest in one unit of the public good. An outcome is then defined as $z = (z_1, ..., z_n)$ with $z_i \in \{0, 1\}$, $\forall i \in I := \{1, \dots, n\}$, where $z_i = 1$ is interpreted as agent i investing his monetary unit into the public good and $z_i = 0$ represents agent i keeping his monetary unit for himself. For notational convenience define $\underline{z} = (0, \dots, 0)$.

Further, all agents $i \in I$ have a valuation $\theta_i \in [0, 1)$ for the public good.[1] Utility of agent i is then given by a quasilinear utility function of the form

$$u_i = 1 - z_i + \theta_i \sum_{j=1}^{n} z_j. \tag{1.1}$$

Valuations θ_i are further assumed to be such that some outcome z exists, which is a strict Pareto improvement over $\underline{z}$ for all agents i, who contribute in z. This assumption ensures that some strict improvement over $\underline{z}$ is possible.[2]

We do not make any specific assumption on whether agents are informed on the valuations of other agents or not. Nash equilibrium is considered only as a first predictor of

[1] Values $\theta_i < 0$ are excluded, since then the public good would be a bad for those agents. If this were the case a mechanism that does not use transfers can never guarantee Pareto improvements. Thus, the mechanism proposed in this paper should only be applied if valuations of the public good of all agents are weakly positive. Values $\theta_i \geq 1$ are excluded for simplicity of notation. Any agent with $\theta_i \geq 1$ has a weakly dominant strategy to contribute the entire endowment to the public good. Thus, there is no need to provide additional incentives to this kind of agents. Therefore, including the possibility of $\theta_i \geq 1$ would not lead to a significant change in any results of the paper, but would complicate notation at several points.

[2] If this were not the case, any Pareto improvement would rely on some agent's contribution, who is indifferent between this Pareto improvement and $\underline{z}$. No mechanism with the desired properties can be asked to provide strict incentives to contribute for this agent in such an environment. Thus, such cases are not considered in the equilibrium analysis.

possible dynamically stable outcomes and the main solution concept is a dynamic adjustment process. Therefore, the results of this paper apply whenever this adjustment process describes behavior reasonably well. This might be the case in environments with complete or incomplete information.

1.2.1 The mechanism

In the Binary Conditional Contribution Mechanism $G^{BCCM} := (M^{BCCM}, g^{BCCM})$ every agent can choose a natural number between 1 and $n+1$. Thus the message space is defined as $M^{BCCM} = \prod_{i=1}^{n} M_i^{BCCM}$, with $M_i^{BCCM} := \{1, 2, \ldots, n+1\}$, $\forall i \in I$. The chosen message is thereby interpreted in the following way: Choosing message $m_i = k$ is like saying "I'm willing to contribute to the public good if at least k agents (including myself) contribute in total." Note that with the messages $m_i = 1$ and $m_i = n+1$ players can decide to contribute in any or no case, respectively.[3]

The outcome selected by the mechanism is the outcome with the highest possible level of contributions such that all those statements are satisfied. Formally, define

$$K(m) := \max\left\{ k \in \{0, 1, ..., n\} \middle| \sum_{i=1}^{n} \mathbb{1}_{(m_i \leq k)} \geq k \right\}. \tag{1.2}$$

The outcome of the mechanism is defined as $g^{BCCM}(m) = z$ with $z_i = 1$ if and only if $m_i \leq K(m)$.[4]

1.2.2 Nash equilibria of the BCCM

The BCCM has multiple Nash equilibria. An example shall demonstrate what properties an outcome must have to be a Nash equilibrium outcome.

Example 1.1. *Consider 5 identical agents with valuation $\theta_i = 0.4 \; \forall \; i \in I$. The trivial Nash equilibrium is given by $m_i = 6$, $\forall \; i \in I$, where no agent contributes to the public good. However, there are more equilibria as e.g. when agents $1, 2$ and 3 choose message $m_i = 3$ and agents 4 and 5 choose $m_i = 6$. In this case the first three agents will contribute to the public good: $z = (1, 1, 1, 0, 0)$. The structure of the mechanism makes*

[3] Since there are only n agents, there can never be $n+1$ contributing agents.

[4] In equation (1.2) $\mathbb{1}_{(m_i \leq k)}$ denotes the indicator function, which is 1 if $m_i \leq k$ and 0 otherwise.

this an equilibrium. Agents 4 or 5 can only change the outcome to $z' = (1, 1, 1, 1, 0)$ *or* $z'' = (1, 1, 1, 0, 1)$ *respectively by unilateral deviation. Neither deviation is beneficial. And the first three agents can only change the outcome to* $\underline{z}$*, which is not beneficial either. Thus, no agent has any incentive do deviate.*

The incentive structure in the example can be generalized. For any outcome there is a message profile that limits the options of agents to the following ones: Agents that currently do not contribute can only alter the outcome by unilaterally contributing themselves, which makes them worse off. Agents that currently contribute can only change the outcome to $\underline{z}$. This indicates that a certain outcome can be implemented as a Nash equilibrium if and only if there is no agent for which the deviation to $\underline{z}$ is profitable.

Theorem 1.1. *z is the outcome of a Nash-equilibrium of the BCCM if and only if* $z \succeq_i \underline{z}, \; \forall\, i \in I$.

Theorem 1.1 predicts equilibria which are Pareto efficient as well as equilibria which may not be Pareto efficient such as those equilibria with outcome $\underline{z}$. Thus, the Nash equilibrium concept does not make a clear prediction as to the equilibrium outcome of the mechanism. Nor does it predict the efficiency of equilibrium outcomes. Therefore, a suitable refinement of the Nash equilibrium concept is needed.

1.3 Unexploitable Better Response Dynamics

As mentioned in the introduction, Better Response Dynamics have been found to describe agents' behavior in repeated public good games rather well (Healy, 2006). Thus, the focus of this section is on Better Response Dynamics as a solution concept. In the following I demonstrate why simple Better Response Dynamics can not be used for the proposed mechanisms. And I motivate a variant of Better Response Dynamics that will be used instead. The intention of this concept is not to perfectly describe subject behavior. The purpose is rather to define a dynamic concept that captures all incentives relevant in the long term. The aim is to correctly predict the set of long term stable outcomes.

Better Response Dynamics assume that a mechanism is played repeatedly by the same agents over a finite or infinite number of periods t. In any period one or more agents are allowed to adjust their message. Agents deviate with positive probability from their current message m_i^t to any message m_i^{t+1} that is a better or best response to m^t. A recurrent class of such a dynamic concept is a set of message profiles, which if ever reached by the dynamics is never left and which contains no smaller set with the same property. If such a recurrent class consists of a single message profile it is called an absorbing state. The equilibrium outcomes of Better Response Dynamics are defined as all outcomes of their recurrent classes.

However, when m_{-i}^t (the message profile containing the messages of all players but i) is fixed, all messages in the BCCM of agent i will lead to only two possible outcomes. This implies that agents will myopically be indifferent between most of their messages. A dynamic adjustment process that only considers myopic better or best response behavior will then have the entire strategy space as its only recurrent class. Thus, simple Better Response Dynamics are not restrictive enough as a solution concept.

I propose to combine the myopic better response condition with a second condition on behavior that is less myopic. Consider the following example.

Example 1.2. *Assume there are 5 identical agents all with type $\theta_i = 0.4$. Assume that currently 4 agents contribute to the public good. The message profile could e.g. be $m^t = (4, 4, 3, 3, 6)$. In this case agents 1 through 4 contribute to the public good. Consider now agent 1. Any message $m_i^{t+1} \in \{1, 2, 3, 4\}$ is a better response for agent i to the message profile m^t. None of those messages would change the outcome if no other agent changes his message at the same time. However, the message $m_1^{t+1} = 3$ gives agent 2 an incentive to deviate to $m_2^{t+2} = 6$ in the following period. Under the new message profile $m^{t+2} = (3, 6, 3, 3, 6)$ only agents 1, 3 and 4 would contribute to the public good making those agents worse and agent 2 better off. The same would be true for the messages $m_1^{t+1} = 2$ and $m_1^{t+1} = 1$. Messages $m_1^{t+1} \in \{1, 2, 3\}$ can thus be exploited by agent 2 in a later period, making agent 2 better off and agent 1 worse off. The special structure of the mechanism makes it possible for agents to prevent this kind of incentives for exploitation without having to free-ride themselves.*

From a strategic perspective the exploitable messages in the example provide other agents with incentives to deviate to less cooperative messages. Thus, not choosing those messages can be interpreted like a second order better response behavior. Agents assume that other agents better respond to the message profile and choose of their own better responses the ones that are strategically optimal. There are more arguments that rationalize this behavior. It is easier, however, to provide those arguments once the term "exploitable" and with it Unexploitable Better Response Dynamics are precisely defined.

Definition 1.2. Given a message profile m and an outcome $g(m) = z$, a deviation from m_i to m'_i is called *exploitable* if there is $m_{-i} \in M_{-i}$ such that $z'(m_{-i}) := g(m'_i, m_{-i}) \prec_i z$ and $z'_i(m_{-i}) > 0$. A message m'_i is called *unexploitable* if it is not *exploitable.*

In the following the assumptions of better responding and unexploitability are combined to one behavioral model.[5]

Definition 1.3. In Unexploitable Better Response Dynamics (UBRD) all agents can adjust their message in every period. Agent i switches in period t to message m_i^t with strictly positive probability if and only if

- m_i^t is a (weak) better response to m^{t-1} and
- m_i^t is unexploitable with respect to $z^{t-1} := g^{BCCM}(m^{t-1})$.

Revisit the example from above with this definition in mind.

Example 1.3. *Assume there are 5 identical agents all with type $\theta_i = 0.4$. Let the current message profile be $m = (6, 6, 6, 6, 6)$. In this case no agent contributes and the outcome is $\underline{z}$. Therefore, a message is exploitable in this case if it makes outcomes possible in which an agent is worse off than in $\underline{z}$. Those messages are only $m_i = 1$ and $m_i = 2$. Both messages are weakly dominated by $m_i = 3$. Thus, when the current outcome is $\underline{z}$ a message is exploitable if and only if it is weakly dominated.*

Therefore, unexploitability can be summarized by two assumptions. First, if agents did not yet coordinate on any Pareto improvements, agents do not send weakly dominated messages. Second, once agents coordinated on a positive level of contributions, they do not choose messages that set incentives for other agents to free-ride on their contribution.

[5]Such a model must further specify whether only one or all agents can change their message in a given period. The latter seems more reasonable for most applications (e.g. international environmental agreements). Thus, I assume in the analysis that all agents can adjust their message every period.

Furthermore, it is not necessary that all agents behave in an unexploitable way. If a large enough subgroup of agents behaves according to UBRD, while the rest of the agents is just better responding, the equilibrium outcomes will be as efficient as if all agents behaved according to UBRD. Consider again an example.

Example 1.4. *Assume there are 5 identical agents all with type $\theta_i = 0.4$. Let the current message profile be $m = (5, 5, 5, 1, 1)$. In this case only agents 1 through 3 send an unexploitable message. Nevertheless, neither of the agents can strictly benefit from any deviation. Although agent 4 and 5's messages are exploitable any attempt to exploit those agents would leave only agents 4 and 5 contributing. Thus, total contribution to the public good would go down by 3. This makes all agents worse off. Thus, in this example it is sufficient if 60% of agents behave according to UBRD to support full cooperation.*

1.3.1 Equilibrium properties of the BCCM under UBRD

Under the stated assumptions agents will learn over time not to choose messages which make them worse off. And they will learn to choose messages that prevent other agents from exploiting their contribution offers. Under the combination of those two assumptions an outcome is stable if and only if it is Pareto optimal and no agent would be equally well or better off in $\underline{z}$. The rest of the paper uses the following definition to simplify notation.

Definition 1.4. z' is a *strict** Pareto improvement over z if z' is a Pareto improvement over z, that is strict for all agents with type $\theta_i \neq 0$. [6]

With this definition I can prove the central result for the binary model.

Theorem 1.5. *An outcome $z \in Z$ is an outcome of some recurrent class of the BCCM under UBRD if and only if it is a Pareto optimal outcome and a strict* Pareto improvement over $\underline{z}$.*

Let me again provide an example to improve the intuition for this result:

[6] Agents who do not profit from the public good ($\theta_i = 0$) can never be strictly better off than in $\underline{z}$. If those agents are excluded this definition of *strict** is not necessary. However, the existence of agents with a valuation of $\theta_i = 0$ makes many mechanisms, which try to force agents to cooperate, to lead to outcomes that are not individually rational. It is thus important to include this case to demonstrate that the BCCM can handle it.

Example 1.5. *Consider a case with 5 identical agents all with type* $\theta_i = 0.4$. *The theorem predicts that all outcomes in which 3,4, or 5 agents contribute to the public good are outcomes of recurrent classes of the BCCM. Those outcomes have in common that they are Pareto efficient in a non-transferable utility setting. Assume for example that the current message profile is* $m = (4,4,4,4,6)$. *Then agents 1 through 4 contribute to the public good, while agent 5 does not. Thus, the outcome is* $z = (1,1,1,1,0)$. *Any deviation of agent 5 would lead to* $z' = (1,1,1,1,1)$ *and would thus not be a better response. For agents 1 through 4 messages* $m_i \in \{5,6\}$ *would lead to the outcome* $\underline{z}$. *They are thus not better responses either. Messages* $m_i \in \{1,2,3\}$ *however make outcomes possible in which the agent has to contribute, but total contribution is less than* 4. *Thus, those messages are exploitable. Therefore, the given message profile is a steady state of UBRD.*[7]

1.4 Non-binary Conditional Contribution Mechanisms

The environment can be generalized to a setting in which contribution is not binary, while keeping the mechanism similar. Assume that every agent can invest any amount between 0 and 1 into the public good. Because it is closer to reality and it keeps the dynamic analysis simpler, I assume a smallest indivisible monetary unit of 0.01.[8]

The BCCM can be adjusted to this environment in a very natural way. However, this natural extension turns out to have equilibria under dynamic considerations, which are not Pareto optimal. Nevertheless, this failure of the natural extension is an important motivation for the more complex message space of the Conditional Contribution Mechanism, which will be introduced afterwards.

1.4.1 The Natural Extension Mechanism

The natural extension of the BCCM will assign every agent i the message space $M_i^{NEM} := \{0, 0.01, \ldots, 0.99, 1\} \times \{0, 0.01, \ldots, n - 0.01, n\}$, where $m_i = (\alpha_i, \beta_i)$ is interpreted as "I am willing to contribute α_i to the public good if total contribution is at least β_i." For the

[7] In this example the other steady states are given by $m' = (5,5,5,5,5)$ and $m'' = (3,3,3,6,6)$ (in any permutation)

[8] This discretization resembles the money structure in most countries. All results in the paper hold with any other finite discretization as well as with different levels of income.

analysis in this section I refer to this mechanism as the Natural Extension Mechanism (NEM). The outcome space is then given by $Z := \{0, 0.01, \ldots, 0.99, 1\}^n$, where z_i is the contribution of agent i to the public good in outcome z. $\underline{z} := (0, \ldots, 0)$ is used as before as the outcome with no contribution to the public good by anyone. The level of contribution selected by the mechanism is again the highest level of total contribution such that all conditions are satisfied. Formally, let $Z^{NEM}(m) \subset Z$ be the set of all outcomes that satisfy all conditions in m. This can be formalized by

$$z \in Z^{NEM}(m) \Leftrightarrow \left(z_i = 0 \text{ or } z_i = \alpha_i \text{ and } \sum_{j=1}^{n} z_j \geq \beta_i \right), \ \forall \, i \in I. \tag{1.3}$$

It is easy to see that $z \in Z^{NEM}(m)$ and $z' \in Z^{NEM}(m)$ imply together $z'' = (\max\{z_1, z'_1\}, \ldots, \max\{z_n, z'_n\}) \in Z^{NEM}(m)$. Thus, the outcome of the mechanism is uniquely defined by

$$g^{NEM}(m) = \text{argmax}_{z \in Z^{NEM}(m)} \sum_{i=1}^{n} z_i. \tag{1.4}$$

The structure of Nash equilibria is similar to the binary case:

Theorem 1.6. *An outcome z is an outcome of a Nash equilibrium of the NEM if and only if $z \succeq_i \underline{z}$, $\forall \, i \in I$.*

Revisit the example

Example 1.6. *Each of five agents has type $\theta_i = 0.4$. Assume $z = (0.5, 0.4, 0.3, 0.2, 0.1)$. Then $z \succ_i \underline{z} \ \forall \, i \in I$. This outcome is the outcome of the Nash equilibrium given by $m_i = (z_i, 1.5)$. This is a Nash equilibrium since no agent can reduce his contribution without the outcome becoming $\underline{z}$. And neither can any agent by changing his message increase any other agent's contribution. Thus, the options for unilateral deviations can be reduced to the same cases as in the binary model.*

Unfortunately, the NEM has undesirable equilibria under UBRD as well. The simplest way to show this is by considering an example.

Example 1.7. *Assume again each of five agents has type $\theta_i = 0.4$. Assume further that in period t all agents sent message $m_i^t = (0.1, 0.5)$ and $z^t = (0.1, 0.1, 0.1, 0.1, 0.1)$.*

Let us find all unexploitable better responses in period $t+1$*. Consider w.l.o.g agent 1. Any message* $m_1' = (\alpha_1, \beta_1)$ *with* $\alpha_1 < 0.1$ *and* $\beta_1 > \alpha_1$ *will lead to* $\underline{z}$ *and is thus not a better response. Any message* $m_1' = (\alpha_1, \beta_1)$ *with* $\alpha_1 < 0.1$ *and* $\beta_1 \leq \alpha_1$ *will lead to* $z = (\alpha_1, 0, 0, 0, 0)$ *and is thus not a better response, either. Any message* $m_1' = (\alpha_1, \beta_1)$ *with* $\alpha_1 > 0.1$ *and* $\beta_1 > 0.4 + \alpha_1$ *will lead to* $\underline{z}$ *and is thus not a better response. Any message* $m_1' = (\alpha_1, \beta_1)$ *with* $\alpha_1 > 0.1$ *and* $\beta_1 \leq 0.4 + \alpha_1$ *will lead to* $z = (\alpha_1, 0.1, 0.1, 0.1, 0.1)$ *and is thus not a better response, either. This leaves only messages with* $\alpha_1 = 0.1$*. However of those messages the ones with* $\beta_1 > 0.5$ *lead to* $\underline{z}$ *and are not a better response and the ones with* $\beta_1 < 0.5$ *are exploitable.* $\beta_1 = 0.3$ *e.g. could lead after deviations of the other agents to* $m_j' = (0.05, 0.3),\ \forall j \in \{2, 3, 4, 5\}$ *to* $z' = (0.1, 0.05, 0.05, 0.05, 0.05)$*. In this outcome agent 1 is worse off than in* z^t *but contributes a strictly positive amount. Thus, his message was exploitable. The only unexploitable better response is thus* $m_1' = (0.1, 0.5)$*. This implies that message profile* m^t *is an absorbing state of UBRD. However,* $z^t = (0.1, 0.1, 0.1, 0.1, 0.1)$ *is not Pareto optimal.*

Agents can in this way get stuck on Pareto improvements over $\underline{z}$ which are not Pareto optimal. Any deviation aiming to make further Pareto improvements possible would make the deviating agent worse off in the next period. And such a deviation is infeasible under a better response behavior.

This problem can be solved by letting agents announce more than one tuple of the form (α_i, β_i). This grants agents a higher flexibility in their strategy giving them the opportunity to explore Pareto improvements with some tuples, while securing the current level of cooperation with one other tuple. As it turns out a message of two such tuples is already enough to solve the issue. Simplicity is a further desirable feature of mechanisms once practical implementations are considered. Thus, the mechanism I propose in the following paragraph lets agents announce exactly two tuples.[9]

[9] Depending on the application different versions of the mechanism are possible. The more tuples agents can send, the more flexible they are. Thus, more tuples could lead to faster convergence. However, more tuples also make the mechanism more complicated. Therefore, a reasonable version for applications might be to let agents announce any amount of tuples they choose between one and some upper bound. This gives agents the simple option of choosing one tuple, while also giving them the option to choose very detailed messages. This mechanism is from the theoretical perspective identical to the version in the paper. The paper version is chosen since it simplifies notation, especially in proofs.

1.4.2 The Conditional Contribution Mechanism

I call this mechanism the Conditional Contribution Mechanism $G^{CCM} := (M^{CCM}, g^{CCM})$. Every agent can announce two tuples $\{(\alpha_i^1, \beta_i^1), (\alpha_i^2, \beta_i^2)\} \in M_i^{CCM} := M_i^{NEM} \times M_i^{NEM}$. The outcome $g^{CCM}(m)$ of the CCM is then defined as in the NEM as the outcome with the highest level of contribution consistent with the messages chosen. Let $Z^{CCM}(m) \subset Z$ be the set of feasible outcomes for a message profile m:

$$z \in Z^{CCM}(m) \Leftrightarrow \; z_i = 0 \text{ or } \left\{ \exists l_i \in \{1,2\} : \; z_i = \alpha_i^{l_i} \text{ and } \sum_{j=1}^{n} z_j \geq \beta_i^{l_i} \right\}, \; \forall \, i \in I \quad (1.5)$$

The outcome of the CCM is then uniquely defined by

$$g^{CCM}(m) = \text{argmax}_{z \in Z^{CCM}(m)} \sum_{i=1}^{n} z_i.^{10} \quad (1.6)$$

The additional tuple in the message has no effect on Nash equilibrium outcomes, since only one of the two announced tuples per agent is responsible for the outcome. Such a mechanism can thus only be found and argued for, when dynamic properties are taken into consideration. The CCM has indeed the desired positive dynamic properties:

Theorem 1.7. *An outcome $z \in Z$ is an outcome of some recurrent class of the CCM under UBRD if and only if it is a Pareto optimal allocation and a strict* Pareto improvement over $\underline{z}$.*

An example shall provide some intuition for this result.

Example 1.8. *Consider the example with 5 agents. Each agent has type $\theta_i = 0.4$. Then in all outcomes of recurrent classes 3 agents contribute their entire endowment. The two other agents can contribute any amount. Take for example the outcome $z = (1, 1, 1, 0.5, 0.5)$. This outcome is supported by the messages $m_i = \{(1, 4), (1, 4)\}$ for $i = 1, 2, 3$ and $m_i = \{(0.5, 4), (0.5, 4)\}$ for $i = 4, 5$. The combination of unexploitability*

[10] The outcome can easily be computed by translating the messages of all agents into step-functions, adding them up and taking the highest fixed point of the resulting function. This makes sure that there is no problem in computation, when n is large.

and better responding behavior makes sure that the outcome cannot be left to another outcome with lower contributions and the unexploitability condition implies further that the outcome cannot be left to any outcome with higher contributions since either agent 4 or 5 would be worse off than in z. Consider for example the message $m'_4 = \{(0.5, 4), (1, 5)\}$. *This deviation in itself does not change the outcome, thus it is a better response. However if agent 5 also switches to* $m'_5 = \{(0.5, 4), (1, 5)\}$, *the outcome would change to* $z' = (1, 1, 1, 1, 1)$. *Since* $u_{4/5}(z) = 2.1 > 2.0 = u_{4/5}(z')$ *the messages* m'_4 *and* m'_5 *are exploitable.*[11]

1.5 Non-linear valuation functions

In this section I drop the assumption that valuations are linear and replace it by a weaker assumption. Consider a finite number n of agents with quasi-linear utility functions $u_i(w_i, w_p) = w_i + f_i(w_p)$, where w_i is the private wealth of agent i and w_p is the total amount of wealth invested into the public good by all agents. The functions f_i are only assumed to be weakly increasing in the level of the public good and may differ across agents.[12] Endowment and outcome space $Z := \{0, 0.01, \dots, 1\}^n$ remain unchanged.[13]

In this setting Pareto optimality will not be enough to ensure that an outcome is part of a recurrent class. Utility gained from the public good increases no longer linearly with the contribution towards the public good. Therefore, there might now be groups of agents who benefit from reducing their own contributions even if all other agents would not contribute anything any more.

In the proofs, I use that the options for deviations of coalitions can be limited to outcomes in which no agent outside the coalition contributes. I call such outcomes enforceable, since coalitions cannot force other agents to contribute. When coalitions' options for deviations are limited to their enforceable outcomes, the equilibrium outcomes of the CCM under UBRD can be captured by the core.

[11] Agents 1 through 3 did not actively exploit the messages of agents 4 and 5 in this example. In some sense these agents exploited each other. However, the important point is that the deviation from z to z' is not desirable for agents 4 and 5.

[12] Note that this includes the cases of agents not profiting at all from the public good, or who get satiated at some level.

[13] A further generalization to different endowments for different agents only complicates notation. The mechanism can easily be adjusted by enhancing the message space and all main results would be unaffected.

Definition 1.8. An outcome $z \in Z$ is enforceable for a coalition $S \subset I$ if $z_i = 0\ \forall i \notin S$. The set of all enforceable outcomes for coalition S shall be denoted Z_S

As in the case of Pareto efficiency I use a standard definition of the core for games without transferable utility as e.g. in (Owen, 1982, p. 293).

Definition 1.9. An outcome $z \in Z$ is in the core if there is no $S \subset I$, $S \neq \emptyset$, and $z' \in Z_S$, such that $z' \succ_i z$, $\forall\, i \in S$.

Since I already demonstrated that Nash equilibrium does not even uniquely predict the outcome in the linear case I skip the static analysis and present only the result under UBRD. As in the previous results there needs to be a strict disincentive for agents to deviate. Since the outcome space is finite the usual core definition does not guarantee this.

I therefore need a definition, which is somewhat stronger than the usual core definition to describe the equilibrium outcomes. Possibilities for deviations under indifference need to be excluded.

Definition 1.10. A core allocation z is *strict** for a subset $S \subset I$ of agents if for any enforceable outcome z' of a coalition S' with $S' \cap S \neq \emptyset$ there exists some agent $i \in S'$ with $z \succ_i z'$.

Definition 1.11. Define the subset $S^C(z) \subset I$ via $i \in S^C(z)$ if and only if $f_i(\sum_{i=1}^n z_i) > 0$ as the set of agents that strictly benefit from the amount of public good in z.

Theorem 1.12. *Assume there exists at least one outcome z that is a core allocation and strict* for $S^C(z)$. Then an outcome z' is an outcome of a recurrent class of the CCM under UBRD if and only if it is a core allocation that is strict* for $S^C(z')$.*

If no such outcome exists the result would be a cycling behavior of the dynamics. It is not obvious that the assumption of existence of such an outcome is satisfied in all relevant cases. However, the existence problem only exists on an infinitesimal level. This is shown, by proving that the mechanism can be adjusted to guarantee existence at arbitrarily low expected costs.[14]

[14] Since costs are arbitrarily low I do not want to argue here who should pay those costs. Note though that in reality costs for setting any such incentives can never be arbitrarily low since the administration costs will be strictly positive. However, the theorem is not meant to fix the problem in applications, but rather to show that the problem is likely to have no effect in real applications at all. Note further that only expected costs can be arbitrarily low as the assumption of a smallest monetary unit makes arbitrarily low payments only possible as lotteries.

In the following theorem let Δ be a mapping from $Z \times I \to \mathbb{R}_+$. The interpretation is that the mapping defines for any agent and any outcome some expected payment $\Delta(z,i) := \delta_{zi}$ that agent i gets payed if outcome z occurs. I write $G + \Delta$ to describe a mechanism G to which the additional payments Δ are added.

Theorem 1.13. *For any environment with weakly increasing valuation functions and for any $\epsilon > 0$ there exists a mapping Δ such that in the game $CCM + \Delta$ there exists a core allocation z, which is strict* for the subset $S^C(z)$. Further, the expected cost of Δ is less than ϵ.*

1.6 Summary and discussion

This paper introduces the class of Conditional Contribution Mechanisms for the provision of public goods. In these mechanisms agents can condition their contribution on the total contribution of all agents. There are efficient as well as inefficient Nash equilibrium outcomes. However, under Unexploitable Better Response Dynamics all equilibrium outcomes turn out to be Pareto efficient, in the non transferable utility sense.

A new concept, Unexploitable Better Response Dynamics, is introduced in the paper to predict the outcomes of the mechanisms. Although the concept is close to the standard concept of Better Response Dynamics and the new unexploitability condition can, besides other arguments, be related to eliminating weakly dominated strategies, there always remains some doubt as to the predictive power of a new solution concept. Therefore, experiments with these mechanisms have to be conducted. A first experiment with the binary environment shows that the BCCM significantly outperforms the VCM in terms of contribution rates and Unexploitable Better Response Dynamics is a good predictor for the stable equilibrium outcomes (Reischmann, 2015b).

Good dynamic equilibrium properties combined with ambiguous Nash equilibrium properties indicate that the mechanism might only be suited for repeated public good problems. However, there are a lot of possibilities to adjust the mechanism for a one-shot game such that the dynamic properties are used. As one example the mechanism could be played five times with the highest contribution in the five trials being used as the outcome. This is close to the way in which the auction mechanism studied by Smith

(1979, 1980) makes coordination possible. Further, agents could be allowed to communicate prior to the one shot game. This form of cheap talk communication was already used successfully to increase contributions in a standard VCM public goods game by Isaac et al. (1985). In the VCM agents have a myopic incentive to lie about the message they intend to send. In the CCM agents do not have such an incentive to lie, since failed coordination makes everyone worse off. Thus, communication should work even better with the CCM. Finding the best way to adjust the mechanism to one shot games is an interesting question for further research.

Everything considered, the class of Conditional Contribution Mechanisms is an important addition to the set of public good mechanisms. It satisfies individual rationality, incentive compatibility, and leads under UBRD to Pareto efficient outcomes in repeated public good environments. Furthermore, in the final analysis the only assumption on valuations is that they are weakly increasing in the level of the public good. Those weak assumptions make the mechanism applicable in a wide variety of public good settings.

Appendix A

General notation: In many proofs I have to show that some outcome z is some sort of equilibrium. In those proofs I need to distinguish between two subsets of agents. The subset of agents who contribute to the public good in z, shall be called $I_1 \subset I$. And the subset of agents who do not contribute to the public good in z shall be called $I_0 \subset I$. If I need a second outcome z' in the proof, those sets will be called I_1' and I_0', respectively.

Proof of Theorem 1.1. Let z be an allocation such that no agent strictly prefers $\underline{z}$ to z and define $k := \sum_{i=1}^{n} z_i$. Then the message profile $m_i = k \; \forall i \in I_1, m_i = n+1 \; \forall i \in I_0$ is a Nash equilibrium with the desired outcome. It is obvious that $g^{BCCM}(m) = z$. In the following I show that m is a Nash equilibrium.

If some agent i in I_1 deviates to a message $m_i' < k$, the outcome does not change. If he changes his message to some $m_i' > k$, the new outcome will be $\underline{z}$. Since no agent strictly prefers $\underline{z}$ to z, this can not make agent i strictly better off. Thus agents in I_1 have no strict incentive to deviate.

If some agent j in I_0 deviates to $m_j' > k+1$, the outcome does not change. If he changes his message to $m_j' \leq k+1$ he will contribute and total contribution will be $k+1$. Since $\theta_j \in [0,1)$ this will make him worse off. Thus also the agents in I_0 have no incentive to deviate and m is indeed a Nash equilibrium.

Let on the other hand z be an outcome such that any agent i strictly prefers $\underline{z}$ to z. Let then m be any message profile leading to the outcome z. By choosing the message $m_i' = n+1$ any outcome that might occur is at least as good for agent i as $\underline{z}$. Thus i has an incentive to deviate. Thus m can not be a Nash equilibrium. □ □

Proof of Theorem 1.5. I prove the theorem in two steps. In step 1, I show that any outcome with the described properties is an outcome of a recurrent class of the dynamics. In step 2, I show that from any other outcome the dynamics reach such a recurrent class with strictly positive probability.

Step1: In the discussion of the environment I assumed that there exists some Pareto improvement z over $\underline{z}$, which is strict for all $i \in I_1$. Such a Pareto improvement is further strict for all agents i with $\theta_i > 0$.

Let z be any such outcome and let $k = \sum_{i=1}^{n} z_i$. Then $m_i = k$ if and only if $i \in I_1$ and $m_i = n + 1$ if and only if $i \in I_0$ is part of a recurrent class of UBRD with outcome z. I prove this by checking that no deviation to a different outcome is compatible with UBRD.

For any agent $i \in I_1$ deviations to any $m_i = k' > k$ will lead to the outcome $\underline{z}$. Since z is a strict Pareto improvement over $\underline{z}$ for those agents this is not a better response. Deviations to any $m_i = k' < k$ make outcomes possible in which i contributes but total contribution is less than k. Thus those strategies are exploitable. Thus no agent in I_1 will change their message according to UBRD. If only agents in I_0 change their messages total contribution can only increase. No agent $i \in I_0$ will choose any $m_i = k' < k + 2$ since then this agent i would contribute. Since $\theta_i \in [0, 1)$ agent i would be worse off. Thus this is not a better response for agent i.

Assume now that after some deviations of agents $i \in I_0$ under UBRD the outcome nevertheless changes from z to z'. Since z was Pareto optimal at least one agent, call him j, is worse off in z' than in z. Since we already noted that no agent in I_1 has any incentive to deviate total contributions are higher in z' than in z. Thus $j \in I_1'$ or agent j could not be worse off in z'. This implies that the messages of agent j that made the change from z to z' possible was exploitable. Thus, j would not have chosen this message under UBRD. And z is indeed the outcome of a recurrent class of the UBRD process.

Step2: Take now any outcome $z \in Z$ which is not Pareto optimal or not a strict Pareto improvement over $\underline{z}$ for all i with $\theta_i > 0$. Then I distinguish two cases. In case 1 z is Pareto optimal but not a strict Pareto improvement over $\underline{z}$ for all i with $\theta_i > 0$. Then there exists some agent i, who contributes, but would be better off by or indifferent to not contributing even if this will lead to $\underline{z}$. Thus for this agent $m_i = n + 1$ is a (weak) better response. Further $m_i = n + 1$ can never be exploitable. If all other contributing agents chose unexploitable messages the switch to $m_i = n + 1$ will lead to the outcome $\underline{z}$. From $\underline{z}$ the dynamics reach any recurrent class with Pareto optimal outcome z, which is a *strict** Pareto improvement over $\underline{z}$, with positive probability. All messages in any such recurrent class are unexploitable better responses, whenever the current outcome is $\underline{z}$.

In case 2 z is not Pareto optimal. Then there exists a Pareto optimal outcome z', which is a Pareto improvement over z. Assume that in z', k' agents will contribute. Then for those agents who contribute in z' but not in z, $m_i = k'$ is an unexploitable better response. Once all those agents play $m_i = k'$, the outcome switches to z'. Thus the dynamics reach z' with positive probability. Now z' is either a Pareto optimum which is a *strict** Pareto improvement over $\underline{z}$, or we are in case 1. □ □

Proof of Theorem 1.6. Let $z := (z_1, \ldots, z_n) \in Z$ be an outcome, such that $z \succeq_i \underline{z} \; \forall \; i \in I$, and define $\bar{\beta} := \sum_{i=1}^{n} z_i$. Then $m_i = (z_i, \bar{\beta})$ is a Nash-equilibrium of the mechanism with outcome z. There are four ways in which any agent i can deviate from this message. He can increase or decrease his proposed contribution. And he can increase or decrease his condition.

Any decrease in the offered contribution will fail to satisfy all other agents conditions and can thus only lead to outcomes, which are worse for agent i, no matter what condition he choses.

Any (weak) increase in the offered contribution will not lead to an increase of other agents' contributions. Thus, such an increase combined with a condition that can be satisfied will only lead to a (weakly) higher contribution by agent i. If the increase in the offered condition is combined with a condition that can not be satisfied the outcome will be $\underline{z}$. In both cases agent i is (weakly) worse off. Thus, no agent has any incentive to deviate and m is a Nash equilibrium.

Let now $z \in Z$ be an outcome such that some agent i strictly prefers $\underline{z}$ to z. Given any message profile m' leading to the outcome z agent i can profitably deviate to $m_i'' = (0, 0)$. This gives him an outcome which is at least as good as $\underline{z}$ and thus strictly better than z. Therefore, there is no message profile that makes z a Nash equilibrium outcome.□ □

Proof of Theorem 1.7. I prove this theorem in two steps. In step 1 I prove that the described outcomes are indeed outcomes of recurrent classes of UBRD. And in step 2 I prove that from any other outcome the dynamics reach one of those recurrent classes with strictly positive probability.

Step1: In the discussion of the environment I assumed that there exists some Pareto improvement z over $\underline{z}$, which is strict for all $i \in I_1$. Take then any Pareto optimal outcome z', which is a Pareto improvement over z. Then z' is a Pareto optimal outcome, which is strict for all $i \in I_1'$. Assume to the contrary that some $i \in I_1'$ were indifferent between z' and $\underline{z}$, then his valuation θ_i must be positive. But then i was either better off in z than in z' if $i \in I_0$, or he was worse off in z than in $\underline{z}$ if $i \in I_1$. Both possibilities lead to a contradiction. Note further that any Pareto improvement z over $\underline{z}$, which is strict for all $i \in I_1$ is further strict for all agents i with $\theta_i > 0$.

Thus, there exists a Pareto optimal outcome $z \in Z$, which is a strict Pareto improvement over $\underline{z}$ for all agents i with $\theta_i > 0$. Let z be such an outcome and define $\bar{\beta} := \sum_{i=1}^{n} z_i$. Then $\alpha_i^1 = \alpha_i^2 = z_i$ and $\beta_i^1 = \beta_i^2 = \bar{\beta}$ is part of a recurrent class of UBRD with outcome z. Assume to the contrary that after deviations of some agents consistent with UBRD the outcome changes from z to some $z' \neq z$. Note that $z' \neq z$ implies in this environment that not all agents are equally well off in z' as in z. Then at least one agent is worse off in z' than in z (otherwise this would be a Pareto improvement over z). If one of the agents who is worse off contributes in z' a strictly positive amount then his message that led to the outcome z' was either exploitable or no better response and he would not have chosen it in UBRD. Thus, all agents, who are worse off in z' than in z, need to contribute zero in z'. Assume to the contrary that in the group of the other agents who are equally well or better off in z' than in z there are some agents who contribute more in z' than in z. Then it would be a Pareto improvement over z if those agents made the contributions as in z', while all other agents made contributions as in z. This cannot be the case since z was Pareto optimal. Thus, all agents contribute weakly less in z' than in z. This implies that total contributions are lower in z' than in z. Then there is one agent in this group whose contribution sank relatively to the contributions in z by the lowest percentage. If this agent is better off in z' than in z he would still be better off in $\underline{z}$ since the valuation of the public good is linear. This contradicts that z was a *strict** Pareto improvement over $\underline{z}$. This yields a contradiction and thus it is not possible that the outcome changes under UBRD once the described message profile is reached.

Step2: Assume now that the current outcome z is not Pareto optimal. Then there exists a Pareto improvement z' over z such that z' is Pareto optimal. Define again $\bar{\beta} := \sum_{i=1}^{n} z_i$ and $\bar{\beta}' := \sum_{i=1}^{n} z_i'$. Then for any agent i the message $\alpha_i^1 = z_i$, $\beta_i^1 = \bar{\beta}$, $\alpha_i^2 = z_i'$, $\beta_i^2 = \bar{\beta}'$ is an unexploitable better response to their current message. If all agents choose this message the outcome will be z'. Thus the dynamics reach this message profile with strictly positive probability. Once it is reached the new outcome is z' and now $\alpha_i^1 = z_i'$,

$\beta_i^1 = \bar{\beta}'$, $\alpha_i^2 = z_i'$, $\beta_i^2 = \bar{\beta}'$ is an unexploitable better response for all agents. Thus from any not Pareto optimal outcome a message profile, like the one in the first part of this proof, is reached with strictly positive probability.

If z' is a strict Pareto improvement over $\underline{z}$ for all agents i with $\theta_i > 0$ the proof is complete. If it is not, then there exists some agent $i \in I_1'$ who is at least as well off in $\underline{z}$ as in z'. For this agent the message $\alpha_i^1 = 0$, $\beta_i^1 = 0$, $\alpha_i^2 = 0$, $\beta_i^2 = 0$ in an unexploitable better response. Thus the dynamics move from any Pareto optimum like z' to $\underline{z}$ with positive probability. From $\underline{z}$ any Pareto optimal allocation, which is a *strict** Pareto improvement over $\underline{z}$, is reached with positive probability in the way described above. □ □

Proof of Theorem 1.12. In the first part of the proof I show that any core outcome z, which is *strict** for $S^C(z)$, is an outcome of recurrent classes of the dynamics.

Let z be an outcome of the mechanism and let z be a core allocation, which is strict* for $S^C(z)$. Define $\bar{\beta} := \sum_{i=1}^n z_i$. Then $\alpha_i^1 = \alpha_i^2 = z_i$ and $\beta_i^1 = \beta_i^2 = \bar{\beta}$ is part of a recurrent class of UBRD with outcome z. Assume to the contrary that after deviations of some agents consistent with UBRD the outcome changes to some $z' \neq z$. Then at least one agent $i \in I_1'$ is worse off in z' than in z (otherwise this would be a coalition improvement over z). Agent i's message, which led to the outcome z', was thus either exploitable or no better response and he would not have chosen it in UBRD.

In the second part of the proof I show that from all other allocations the dynamics move with strictly positive probability to a core allocation, which is *strict** for $S^C(z)$.

Assume that the dynamics are in a state with some outcome z, which is not Pareto optimal and let z' be any Pareto optimal allocation, which is a Pareto improvement over z. Define $\bar{\beta} := \sum_{i=1}^n z_i$ and $\bar{\beta}' := \sum_{i=1}^n z_i'$. Then the message $(z_i, \bar{\beta}), (z_i', \bar{\beta}')$ is an unexploitable better response for any agent i. Thus the dynamics move with strictly positive probability from z to any such z'.

I can thus assume that the dynamics are in a state with some outcome z, which is Pareto optimal, but not a core outcome that is strict for $S^C(z)$. Then there exists a coalition S and an outcome $z' \in Z_S$ such that all agents $i \in I_1'$ are at least as well off in z' than in z. This implies that $\bar{\beta}' := \sum_{i=1}^n z_i' < \bar{\beta} := \sum_{i=1}^n z_i$ or this would be a Pareto improvement. Then in a first step the messages $(z_i, \bar{\beta}), (z_i', \bar{\beta}')$ are unexploitable better responses for every agent $i \in I_1'$. Once all agents $i \in I_1'$ switched to those messages, the messages $(z_i', \bar{\beta}'), (z_i', \bar{\beta}')$ and $(z_i, \bar{\beta}), (z_i, \bar{\beta})$ are both unexploitable better responses for those agents, since the current outcome is still z. But if now simultaneously one agent chooses $(z_i', \bar{\beta}'), (z_i', \bar{\beta}')$ and another one chooses $(z_j, \bar{\beta}), (z_j, \bar{\beta})$, then contribution breaks down entirely and the outcome will be $\underline{z}$. From $\underline{z}$ any core allocation, which is a Pareto improvement over $\underline{z}$ and *strict** for $S^C(z)$ will be reached with strictly positive probability in the way described above. □ □

Proof of theorem 1.13. I prove this theorem in two steps. In step 1, I show that it is possible to design arbitrarily cheap incentive schemes, such that no agent is indifferent between any two outcomes. In step 2, I show that this leads to the existence of a core outcome in the given environment. Finally, when every agent has a strict preference between any two outcomes then any core outcome is *strict** for all subsets of agents. Thus, there exists a core outcome z, which is strict for $S^C(z)$.

Step 1: Let $\epsilon > 0$. Define $\epsilon' := \min_{i \in I} \min_{z,z' \in Z: u_i(z) \neq u_i(z')} |u_i(z) - u_i(z')|$ as the smallest positive difference in utility between any two outcomes for any agent. Let $N_Z := \#Z$ be the number of possible outcomes and let $r : Z \to \{1, ..., N_Z\}$ be any bijective mapping, which satisfies $\sum_{i=1}^n z_i > \sum_{i=1}^n z_i' \Rightarrow r(z) > r(z')$. Define the mapping $\Delta_{zi} = \frac{r(z) \min(\epsilon, \epsilon')}{2nN_Z}$ $\forall\, i \in I$. Total cost of this mapping can be estimated in the following way:

$$\sum_i^n \Delta_{zi} = \sum_i^n \frac{r(z) \min(\epsilon, \epsilon')}{2nN_Z} \leq \sum_i^n \frac{N_Z \min(\epsilon, \epsilon')}{2nN_Z} \leq \frac{n \min(\epsilon, \epsilon')}{2n} \leq \frac{\epsilon}{2} \tag{1.7}$$

Thus the mapping has total cost of at most $\frac{\epsilon}{2}$. Assume now to the contrary that some agent i is indifferent between any two outcomes z and z' under the mechanism with the incentive scheme Δ. This indifference implies:

$$u_i(z) + \Delta_{zi} = u_i(z') + \Delta_{z'i} \Leftrightarrow u_i(z) - u_i(z') = \Delta_{z'i} - \Delta_{zi} \tag{1.8}$$

The absolute value of the left-hand side of this equation is either equal to zero or weakly bigger than ϵ'. However, since $r(z) \neq r(z')$ the absolute value of the right-hand side is strictly bigger than zero and strictly smaller than ϵ'. This leads to a contradiction. Therefore, adding the incentive scheme Δ leads to a mechanism in which no agent is indifferent between any two outcomes.

Step 2: I prove this step by induction over the number of agents in the economy. For the beginning assume there are $n = 1$ agents. Then existence of a core outcome is equivalent to the existence of an outcome which gives the agent maximal utility. Since our state space is finite this is trivial. Thus, one may assume that for an economy with $n = k$ agents there exists a core outcome. Let's now look at an economy with $n = k+1$ agents. Call the coalition of agents 1 through k in this economy C. Then by assumption there is an outcome z, with $z_{k+1} = 0$, from which no subcoalition of C can improve. I call this a core outcome in the coalition C. Let z' be the Pareto optimal Pareto improvement over z, in which agent $k+1$ gets the highest utility. Then no subcoalition of C can improve on z'. Otherwise z could not have been a core outcome in coalition C. Assume to the contrary a coalition C' including agent $k + 1$ can improve from z' to an outcome z''. Then total contributions are less in z'' than in z' or this would be a further Pareto improvement. Then $z''' := (\max\{z_1, z_1''\}, ..., \max\{z_k, z_k''\}, z_{k+1}'')$ is a Pareto improvement over z in which agent $k + 1$ is better off than in z'' (since $\sum_{i=1}^{n} z_i > \sum_{i=1}^{n} z_i' \Rightarrow r(z) > r(z'))$ and thus better off than in z'. This contradicts the assumptions on z'. Thus, no coalition can improve on z' and therefore z' is in the core.□ □

Chapter 2

Conditional vs. Voluntary Contribution Mechanism An Experimental Study *

Abstract

The Conditional Contribution Mechanism for public good provision gives all agents the possibility to condition their contribution on the total level of contribution provided by all agents. In this experimental study the mechanism's performance is compared to the performance of the Voluntary Contribution Mechanism. In an environment with binary contribution and linear valuations subjects play the mechanisms in a repeated setting. The mechanisms are compared in one case of complete information and homogeneous valuations and in a second case with heterogeneous valuations and incomplete information. In both cases a significantly higher contribution rate can be observed when the Conditional Contribution Mechanism is used.

*I thank the participants of the IWGTS at the University of Sao Paulo. I also thank Jörg Oechssler, Peter Dürsch, Florian Kauffeld, Hannes Koppel, Christoph Brunner, Jonas Hedlund and Christoph Vanberg as well as all participants of the seminars at Heidelberg University for helpful comments.

2.1 Introduction

Numerous attempts have been made to solve the free-rider problem in public good environments. While there are many complex mechanisms that have good theoretical properties, it is exactly this complexity that makes them difficult to apply in practical applications. However, the simple mechanisms which are mostly used, like e.g. the Voluntary Contribution Mechanism (VCM), do not have good theoretical properties and suffer, at least to some extent, from the free-rider problem. With the recent development of the Binary Conditional Contribution Mechanism (BCCM) there is a new simple candidate to solve the free-rider problem (Reischmann, 2015a). This paper presents the first experimental evidence on the performance of the BCCM.

This special case of the class of Conditional Contribution Mechanisms (CCMs) is applicable in binary contribution environments. It extends the message space of the VCM {"Contribute", "Don't Contribute"} by conditional contribution offers of the form "Contribute only if at least k other agents contribute as well". The mechanism is played simultaneously by all agents. When there are multiple outcomes that satisfy all conditions, the mechanism selects of these outcomes the one with the largest amount of total contributions.

The CCMs are designed with a focus on dynamic properties. Thus, the static equilibrium properties are not very impressive. The CCMs have many efficient, but also many inefficient Nash equilibria. However, Reischmann (2015a) applies a variant of Better Response Dynamics under which all outcomes of dynamic steady states are Pareto efficient. Since Healy (2006) shows that Better Response Dynamics describe subject behavior in public good games rather well, the BCCM is well suited for repeated public good environments.

The aim of this experimental study is to evaluate whether the BCCM is a suitable candidate to solve the free-rider problem. For this sake, I compare the BCCM with the VCM. The VCM is chosen as a comparison since it is, besides the Provision Point Mechanism (PPM), the only mechanism that is regularly applied in practical applications. Further the PPM is better suited for step-level public goods, which are not the focus of this study. Thus, the VCM is still the most important benchmark to beat.

Both the VCM and the BCCM are tested in two repeated public good environments, one with complete and one with incomplete information. In both treatments I find the effect, that the BCCM produces significantly higher contribution rates than the VCM. As expected from the theoretical analysis the difference in contribution rates is mainly found in the last 10 periods. This result supports the theoretical prediction that the BCCM sets better dynamic incentives in repeated public good environments than the VCM.

2.1.1 Related literature

My work mostly relates to two kinds of literature, first experiments comparing the performance of two or more public good mechanisms and second experimental studies on behavior in public good mechanisms in general, or the Voluntary Contribution Mechanism in particular.

Smith (1979, 1980) compares his auction mechanism to the VCM and a quasi-free-rider mechanism. All three mechanisms have in addition an unanimity rule. If an outcome is not unanimously accepted no contribution will be made to the public good. Smith finds that the auction mechanism supplies significantly higher levels of the public good as the free rider quantity. However, if the cases when unanimity fails are taken into account the auction mechanism does not perform significantly better than the alternative mechanisms. Banks et al. (1988) continue the investigation of the auction mechanism and compare it to the VCM. They compare both mechanisms with and without unanimity. They find that the auction mechanism is more efficient than voluntary contribution. Unanimity seems to lower contributions overall.

There are multiple studies comparing the VCM with a Provision Point Mechanism (PPM). Rondeau et al. (1999) show that under specific conditions the PPM can be demand revealing. Building on this Rondeau et al. (2005) find that the PPM leads in a lab and a field experiment to a higher willingness to contribute to the public good than the VCM. Rose et al. (2002) further study the PPM in a field experiment on green energy and find that contribution rates outperform previous studies that used the VCM. However, every contribution in the VCM can definitely be used to finance a certain level of the public good. Contributions under the PPM might be lost for the public good if the threshold for provision was chosen too high.

Further mechanisms that have recently been tested experimentally are auction and lottery mechanisms. Schram and Onderstal (2009) compare a first-price winner-pay auction, a first-price all-pay auction and a lottery. They find that out of those three mechanisms the all-pay auction leads to significantly higher contributions. Morgan and Sefton (2000) present an experiment in which a lottery leads to higher contribution to a public good than the VCM. They further find that higher price money leads to a more effective mechanism. Contrary to the findings of Schram and Onderstal (2009), Corazzini et al. (2010) show an experiment in which a lottery outperforms an all-pay auction. Still in their experiment both mechanisms fare better than the VCM.

Behavior in public good environments under the Voluntary Contribution Mechanism is very well understood. Early experiments find consistently that contribution rates are around half-way between the efficient and the free-rider quantity in one-shot games. Under repeated interaction these contributions decline over time. See Ledyard (1994) for a survey on this branch of the literature. The more recent experiment by Burger and Kolstad (2009) covers VCM treatments with binary contributions and they find results in the same spirit, medium contribution rates in the first period and a decline of contributions over time.

By now economic theory can explain these findings that contradict the strong free-rider hypothesis. One explanation is given by social preferences, as e.g. the model of inequality aversion by Fehr and Schmidt (1999). Another explanation are preferences for conditional cooperation as found by Fischbacher et al. (2001). In combination with the finding of Healy (2006) that agents better respond in public good environments this explains positive contributions in the first period as well as the decline over time.

The rest of the paper is structured as follows. Section 2.2 introduces the two mechanisms that will be covered in the experiment. In section 2.3, there is a short theoretical analysis of the equilibrium properties of those mechanisms. Section 2.4 covers the description of the experimental setup, while section 2.5 presents the results. Finally, section 2.6 gives a short summary and discussion of the paper. Translations of written instructions and test questions handed out to subjects in the experiment can be found in Appendix B and C.

2.2 Environment and mechanisms

In this experimental study I compare two different mechanisms. Both mechanisms are tested in an environment with five agents. Those agents play one of the mechanisms as a stage game repeated over 20 periods. In every period the agents are endowed with 10 points.[1] Those points can be invested in a group project or be kept in the private account. The points can not be divided between the two options, so contribution is binary. In the following section an outcome is going to be described by $z = (z_1, z_2, z_3, z_4, z_5)$. $z_i = 0$ denotes that agent i does not invest his points into the project and $z_i = 10$ implies that he does invest his points into the project.

The two mechanisms are compared in two different cases. One case with complete information and homogeneous valuations and one with incomplete information and heterogeneous valuations. Comparisons between the two cases with the same mechanism are not the focus of this study.

In the complete information case every agent knows all players' valuations for the public good and valuations are homogeneous with $\theta_i = 0.6$. In the incomplete information case agents only know their own valuation and all agents have a valuation of $\theta_i = 0$ with a probability of 20% and a valuation of $\theta_i = 0.6$ with probability of 80%. Thus, heterogeneous valuations are possible. The first type of agents, who do not benefit from the public good, are called type 1 agents. And agents, who do benefit, are called type 2 agents. The draws of valuations for group members are independent. Every draw is used though for one group with each mechanism to ensure comparability. Given their valuation agents have the following payoff function:

$$\Pi_i = 10 - z_i + \theta_i \sum_{j=1}^{5} z_j \tag{2.1}$$

In the Binary Conditional Contribution Mechanism (BCCM) agents can condition their contribution on a total level of contribution provided by all agents. The message space is given by $M_i = \{0, 1, 2, 3, 4, 5\}$. Message $m_i = k$ can be interpreted as saying "I'm willing to contribute to the public good if at least k other agents contribute as well." Message $m_i = 0$ is equivalent to contributing in any case. And the message $m_i = 5$

[1] 10 points are chosen to ensure that the number of points earned in each period is a natural number in all cases.

is equivalent to contributing in no case.[2] An offer for conditional contribution can be satisfied it two ways. Either the agent does not have to contribute. Or his condition for contribution is satisfied. This implies that for a given message profile m there might be more than one outcome z that satisfies all conditional contribution offers. Of all those outcomes the BCCM selects the one with the highest total level of contribution. This can be formalized by using the following help variable:

$$K(m) := \max \left\{ k \in \{0, 1, ..., n\} \middle| \sum_{i=1}^{n} \mathbb{1}_{(m_i < k)} \geq k \right\}, \tag{2.2}$$

where $\mathbb{1}_{(m_i<k)}$ denotes the indicator function, which is 1 if $m_i < k$ and 0 otherwise. With this variable, the outcome of the mechanism can be defined as $g^{BCCM}(m) = z$ with $z_i = 10$ if and only if $m_i < K(m)$.

Example 2.1. *Consider the following examples. If all agents choose $m_i = 5$, no agent contributes to the public good. If all agents choose $m_i = 4$, there are two outcomes that satisfy all conditions. In the outcome $z = (10, 10, 10, 10, 10)$ all agents contribute to the public good. And in the outcome $z = (0, 0, 0, 0, 0)$ no agent contributes to the public good. Therefore, the BCCM selects the first outcome and all agents will have to contribute. Similarly if e.g. $m_1 = m_2 = m_3 = 2$, $m_4 = 4$, $m_5 = 5$, then agents 1, 2 and 3 will contribute to the public good.*

The mechanism that I use as a benchmark is a standard binary Voluntary Contribution Mechanism (VCM). In this mechanism agents have only two options. They can contribute in any case or free-ride in any case.

2.3 Theoretical predictions

For an extensive analysis of the theoretical predictions of the general Conditional Contribution Mechanisms I refer to the companion paper (Reischmann, 2015a). Here I analyze the predictions for the specific versions of the mechanisms used in the experiments.

Two different solution concepts will be considered. The first one is Nash equilibrium, since it is the most standard concept and it provides intuition about what might be

[2]Since there are only 5 agents in total, there can never be more than four other contributing agents.

stable outcomes of the mechanisms. The second one is Unexploitable Better Response Dynamics. This is a variant of Better Response Dynamics which is developed and motivated in the companion paper mentioned above.[3] The definition is as follows:

Definition 2.1. Given a message profile m and an outcome $g(m) = z$, a deviation from m_i to m'_i is called *exploitable* if there exists $m_{-i} \in M_{-i}$ such that $z'(m_{-i}) := g(m'_i, m_{-i}) \prec_i z$ and $z'_i(m_{-i}) > 0$. A message m'_i is called *unexploitable*, if it is not *exploitable*.

Definition 2.2. In Unexploitable Better Response Dynamics (UBRD) all agents can adjust their message in every period. Agent i switches in period t to message m_i^t with strictly positive probability if and only if

- m_i^t is a (weak) better response to m^{t-1} and
- m_i^t is unexploitable with respect to $z^{t-1} := g(m^{t-1})$.

Summarizing the motivation given in Reischmann (2015a), UBRD makes the following two assumptions on long term incentives. First, in the long term agents do not choose messages that make them worse off immediately. This is captured by the better response condition. Second, agents do not choose messages that make outcomes possible, in which the agent has to contribute to the public good, but is worse off than in the current outcome. This is captured by the unexploitability condition.

Since this is an experimental study, the experimental results will present a good opportunity to evaluate the validity of this concept. Thus, the discussion whether UBRD is a reasonable solution concept for this mechanism is postponed to section 2.5, where the experimental results are discussed.

2.3.1 Voluntary Contribution Mechanism

The Voluntary Contribution Mechanism gives every agent the choice whether he wants to contribute to the public good. And no agent has any influence over any other agent's contribution. Disregarding social preferences, it is easy to see, and well known in the

[3] UBRD is only supposed to capture all relevant long term incentives of the Conditional Contribution Mechanisms. Thus, the concept wants to make a good prediction about what outcomes occur in dynamically stable states of the CCM. The concept is not intended to describe agents short term behavior in detail. Nor is it intended to be applicable to other mechanisms.

literature, that free-riding is a dominant strategy here. This is true for all agents with a valuation $\theta_i < 1$. The straight forward Nash prediction, not taking possible social preferences into account, is thus that all agents will free-ride. Since free-riding is a strictly dominant strategy any refinement of Better Response Dynamics will also predict this outcome as a unique steady state.

However, it is equally well known that this theoretical prediction is seldom to never observed in experiments. Indeed the general observation is a contribution rate of about 40-60% of the efficient level in the first period. If the public good game is played repeatedly, as it is in this study, the typical experimental finding is that contribution rates decline over time. Social preferences (Fehr and Schmidt, 1999) as well as preferences for conditional cooperation (Fischbacher et al., 2001) in combination with a better responding behavior (Healy, 2006) explain these findings well. Since those findings are very persistent (see Ledyard (1994) for a survey of the early findings and Burger and Kolstad (2009) for a recent example with the binary VCM), this is also what I expect to find in this experiment.

2.3.2 Binary Conditional Contribution Mechanism

Reischmann (2015a) proves that any outcome of the BCCM is the outcome of a Nash equilibrium if and only if it is a weak Pareto improvement over $\underline{z} := (0, 0, 0, 0, 0)$. This is easy to see for the specific case considered in the experiment. Consider the outcome $z = (10, 10, 10, 0, 0)$ in the complete information treatment ($\theta_i = 0.6 \ \forall \ i = 1, \ldots, 5$). This is a Pareto improvement over $\underline{z}$. One Nash equilibrium that leads to this outcome is given by $m_1 = m_2 = m_3 = 2$, $m_4 = m_5 = 5$. All other Pareto improvements z over $\underline{z}$ are supported as Nash equilibrium in similar fashion. Agents who contribute in z condition their contribution on the total level of contribution in z and all other agents choose to contribute in no case.

Thus, the Nash prediction in the homogeneous case would be that either none, two, three, four, or all five subjects contribute in any group. This, of course, is no useful prediction since it only excludes outcomes in which one subject contributes alone. Unexploitable Better Response Dynamics, however, predict convergence to either an outcome in which all five subjects contribute ($m = (4, 4, 4, 4, 4)$), or an outcome in which four agents contribute ($m = (3,3,3,3,5)$, in any permutation). Note that these are exactly the outcomes which are Pareto efficient in a non-transferable utility framework.

The formal proof that the stable outcomes of the BCCM under UBRD coincide with the Pareto efficient allocations, which are Pareto improvements over $\underline{z}$, can again be found in the companion paper. Here I provide some intuition with another example.

Example 2.2. *Assume that the current message profile is* $m = (3,3,3,3,5)$. *In this case only agent 5 does not contribute to the public good. Thus, the outcome is* $z = (10,10,10,10,0)$. *Any deviation of agent 5 will lead to an outcome in which he has to contribute to the public good. This would not be a better response. If any one of agents 1 through 4 switches to a message* $m_i \in \{4,5\}$ *the outcome would be* $\underline{z}$. *Those messages are not better responses either. Any message* $m_i \in \{0,1,2\}$ *violates the unexploitability condition since it makes outcomes possible in which the agent must contribute but is worse off than in the current outcome. Take e.g.* $m_1 = 2$. *This makes the message profile* $m = (2,2,2,5,5)$ *possible. In this profile agent* 1 *has to contribute and total contributions are lower than in the other outcome. Thus, the message profile* $m = (3,3,3,3,5)$ *is one steady state of UBRD.*

The next example demonstrates why outcomes that are not Pareto efficient can not be steady states of UBRD.

Example 2.3. *Assume that the current message profile is* $m = (2,2,2,5,5)$. *In this case agents 1,2 and 3 contribute to the public good. Thus, the outcome is* $z = (10,10,10,0,0)$. *No agent can directly benefit from any deviation. Thus, the message profile* m *is a Nash equilibrium. However, agents 4 and 5 can deviate to the message* $m_i = 4$. *One such deviation does not change the outcome and a unilateral deviation is thus a weak better response. Further, agents 4 and 5 will only have to contribute to the public good if the outcome will be* $z' = (10,10,10,10,10)$. *Agents 4 and 5 are both better off in* z' *than in* z. *Thus, the message* $m_i = 4$ *is unexploitable. However, if both agents 4 and 5 switch to* $m_i = 4$ *the outcome will indeed be* z'. *Therefore,* m *is not a steady state of UBRD.*

The equilibria in the incomplete information treatment mirror the results of the complete information case. Since the dynamics only consider the heterogeneity part of the incomplete information treatment this is not surprising. Thus, in this treatment either all or all but one type 2 agents are predicted to contribute. Still, all outcomes of steady states are Pareto efficient.

2.4 Experimental design

The experiments were conducted at the Alfred-Weber Institute of Heidelberg University. The subject pool used for recruiting consists mainly of students. In each session 10, 15 or 20 subjects participated in groups of 5. In total 195 subjects took part in the experiments. Seven groups played the VCM with complete information and eight groups played the BCCM with complete information. In the incomplete information treatment each mechanism was played by 12 groups. Sessions lasted between 45 minutes and one hour.

When the subjects entered the lab they were randomly allocated to their seats by drawing numbered cards. Every subject was then handed one set of instructions and test questions. English translations of the instructions and test questions can be found in Appendix B and C. Once all subjects answered the test questions correctly and there were no more questions a computer program written in z-Tree (Fischbacher, 2007) was started. The program randomly matched subjects in groups of 5. Groups stayed the same over all 20 periods. Every group played only one mechanism and only one information treatment. In the incomplete information treatment the random draw of types was performed by the program at the beginning of period one.

After the last period there was a short questionnaire asking for personal characteristics such as gender and previous knowledge of game theory. Afterwards subjects were called by seat number to receive their payoff in private. In every period subjects could earn between 6 and 30 points. Points of all periods were added up. Subjects were payed 1€ for every 40 points. Type 1 subjects in the incomplete information treatment received an additional 5€ to compensate them for the lower earning possibilities. Average earnings per subject were 11.55€.

2.5 Experimental results

This study intends to answer two questions. First, is the BCCM suited to improve contribution rates to public goods compared to the VCM? Second, is the model of UBRD suited to predict long term stable outcomes of Conditional Contribution Mechanisms?

2.5.1 Contribution rates

Whether or not the BCCM can increase contributions to the public good significantly compared to the VCM is the primary question. Therefore, I compare total contributions in groups under the BCCM to those contributions under the VCM using the Wilcoxon-Rank-Sum Test. In each group I take the average of total contributions over a certain number of periods. First, I consider all periods to get an impression of the total effect. Second, I only consider the last 10 periods, to get an impression of the long term effect, once a certain level of convergence has taken place including the endgame effect. Third, I consider periods 9 to 18. This choice makes it possible to look at the long term effects excluding the end game effect.

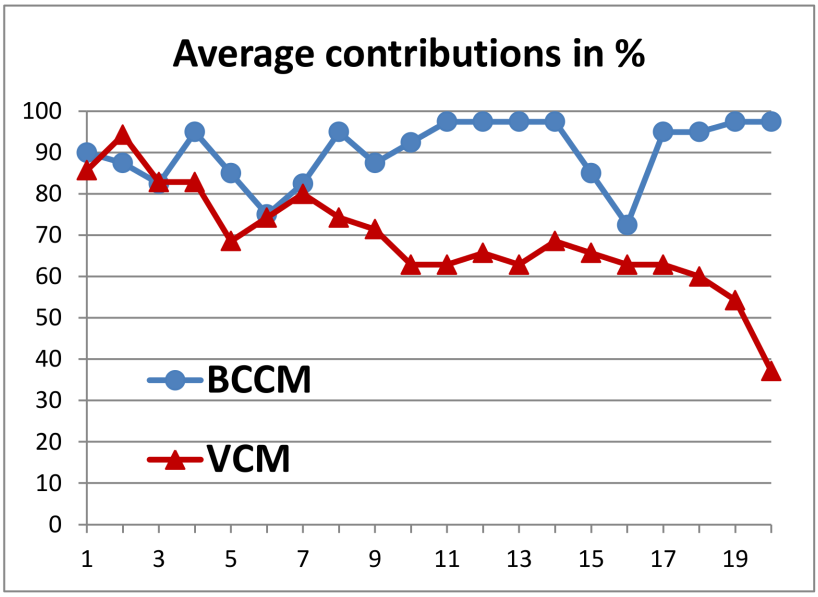

Figure 2.1: Comparison of average contributions over all groups in the complete information treatment.

Average contributions per period over all groups with complete information are displayed in figure 2.1. The figure makes the following immediate observations possible. First, the contribution rate in the first period under the VCM is surprisingly high. The reason for this is probably the binary contribution environment in combination with the rather small group size of 5. Second, contributions in the VCM decline over time as expected. Third, contribution rates in the BCCM are similar to the VCM in early periods but much higher in the later periods. Fourth, the BCCM does, in this treatment, not suffer from any endgame effect. All these observations support the theoretical prediction that the BCCM has better dynamic properties than the VCM. In fact the BCCM leads already to

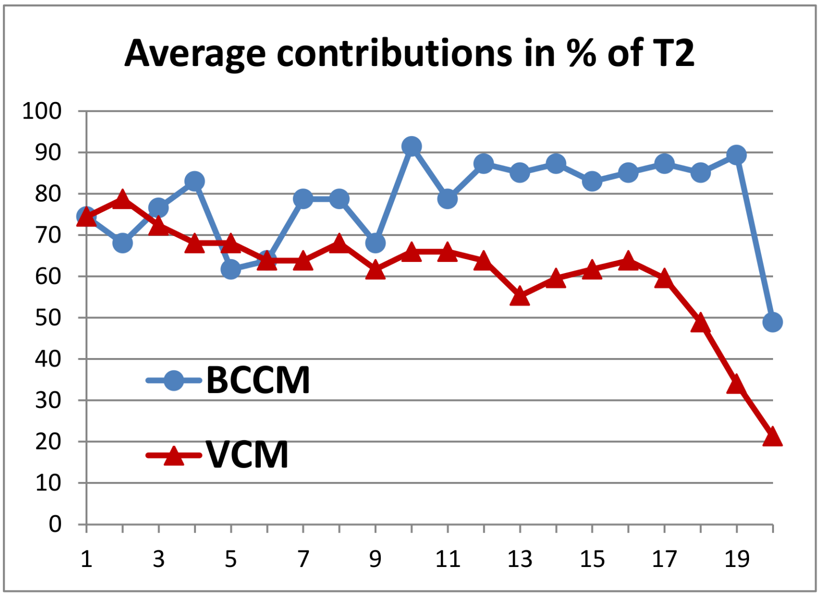

FIGURE 2.2: Comparison of average contributions over all groups in the incomplete information treatment.

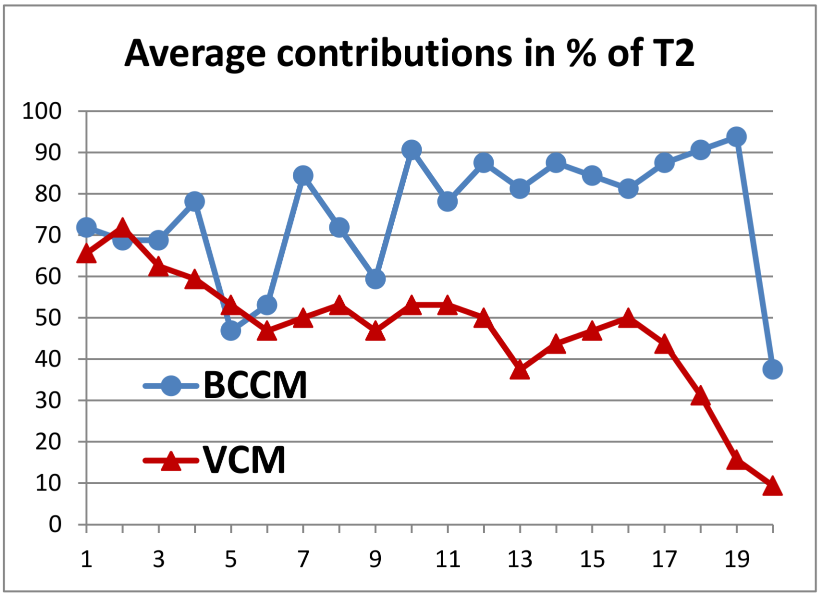

FIGURE 2.3: Comparison of average contributions over all groups with at least one type 1 agent.

significantly higher contributions when all periods are taken into account ($p = 0.0425$). When only the last 10 periods are considered the effect is highly significant ($p = 0.0080$). And when I exclude the endgame effect (periods 9 to 18) the results are still significant ($p = 0.0388$).

In the incomplete information treatment type 1 agents have a dominant strategy to free-ride. Besides a few mistakes in period 1 and one mistake in period 2 all subjects also chose this strategy. Therefore, contribution rates in the incomplete information treatments are always compared in terms of average contributions of type 2 agents. Average contributions of type 2 agents per period over all groups with incomplete information are displayed in figure 2.2. The observations from this figure differ from the complete information case in only one way. Under incomplete information the BCCM suffers from a severe endgame effect. There are two reasons for this. Some agents harm themselves by deviating because they try to free-ride in a coordinated equilibrium resulting in a complete breakdown of contributions. My only explanation for this behavior is that some agents make mistakes, because of the somewhat higher complexity of the incomplete information treatment. The second reason is that some groups reach Pareto efficient outcomes, but no stable equilibrium by period 20. This creates incentives for individual agents to deviate. However, more than one deviation usually leads again to a complete breakdown. This makes the endgame effect even bigger than in the VCM treatments. This second effect might vanish when more periods are played, which gives subjects more time to converge to stable equilibria. Besides this point, however, the results are very similar. The increase in contributions when all periods are considered is at least weakly significant ($p = 0.0602$). For the last 10 periods results are again significant at the 1% level with a p-value of $p = 0.0078$. And when the last two periods are excluded the increase is significant at the 5% level ($p = 0.0199$).

Figure 2.3 shows the average contribution rates when in the incomplete information treatments only those groups are considered that contain at least one type 1 agent. This leads to lower contribution rates under both mechanisms in the first half of the experiment. However, groups playing the BCCM manage to achieve the same high contribution rates in periods 10 to 19. Groups playing the VCM however can not stop the decline of contributions. This observation strengthens the impression that the BCCM robustly reaches high long term contribution rates even in settings in which coordination in the early periods is difficult.

Result 2.3. Under complete as well as under incomplete information the Binary Conditional Contribution Mechanism leads to higher contribution rates than the Voluntary Contribution Mechanism.

The theoretical analysis further suggests that the BCCM should be able to reach stable equilibria with high contribution levels. If this is true it should not be possible to find a decrease in contributions over time as has repeatedly been shown for the VCM. In fact since failed coordination in the early periods might lead to low contribution rates in those periods the BCCM might lead to an increase of contributions over time. Therefore a Wilcoxon Matched-Pairs Signed Rank Test is used to compare average contribution rates over the first 10 periods with those over the last 10 periods.

In both treatments with the Voluntary Contribution Mechanism contributions in the first 10 periods are significantly higher than in the last 10 periods. The p-values are given by $p = 0.0343$ for complete information and $p = 0.0022$ for incomplete information.

In the BCCM treatments on the other hand I observe higher contribution rates in the last 10 than in the first 10 periods under complete information with a p-value of $p = 0.0193$. And under incomplete information the hypothesis that contribution rates in the first and last 10 periods are equal can not be rejected ($p = 0.3668$).

Result 2.4. In treatments with the Voluntary Contribution Mechanism contributions decrease over time. In treatments with the Binary Conditional Contribution Mechanism this is not the case. Under complete information contribution rates are even increasing.

One typical goal of the implementation problem is that the designed mechanism should lead to Pareto efficient outcomes. Whether the BCCM leads to Pareto efficient outcomes can only be answered qualitatively. In the case of complete information when all periods are considered 91.88% of outcomes are Pareto efficient. When only the last 10 periods are taken into account 96.25% of outcomes are efficient. And in the last 4 periods every single outcome is Pareto efficient. Note again that Pareto efficiency is considered without the possibility of transfer payments. Thus, an outcome is Pareto efficient if four or five agents contribute to the public good.

While the theoretical prediction of Pareto efficient outcomes fits the data well in the complete information case the situation differs under incomplete information. In those treatments 75.42% of all outcomes under the BCCM are Pareto efficient. This number increases slightly to 80.83% in the last 10 periods, but decreses again to 75% in the last 4 periods, because of the endgame effect under incomplete information.

Result 2.5. Under complete information the Binary Conditional Contribution Mechanism converges to Pareto efficient outcomes. Under incomplete information about 3 out of 4 outcomes are Pareto efficient.

2.5.2 Unexploitable Better Response Dynamics

Finally, I am interested in the model of Unexploitable Better Response Dynamics itself. How well does the model fit the data for the Binary Conditional Contribution Mechanism?

Note first that the model of better responding agents fits the data pretty well. In the complete information treatment about 93% of messages sent are better responses. In the incomplete information treatment the value is even a little bit higher at 96%, both times high enough to claim that a better responding behavior describes the observations reasonably well. However, only around half of all messages are also unexploitable better responses in the two treatments (41% under complete and 53% under incomplete information).

There is no support for a theory that agents learn to choose unexploitable messages over time under incomplete information (52% of messages are unexploitable better responses in the last ten and 53% in the last 5 periods). And only weak support for a learning towards unexploitablity under complete information (35% in the last ten and 47% in the last 5 periods).

However, Unexploitable Better Response Dynamics is only intended as a concept that predicts long term stable outcomes. As such UBRD predicts that the long term stable outcomes are the Pareto efficient outcomes. If the dynamics are considered to have converged to a stable outcome if at least four out of the last five outcomes are identical then, 14 out of 20 groups converge to an outcome. Of those 14 outcomes all 14 are Pareto efficient. This supports the conclusion that UBRD predicts the dynamically stable outcomes of the BCCM correctly. In comparison, under the definition of convergence from above, 8 out of 19 groups under the VCM reach a stable outcome. Of those 8 outcomes 4 are Pareto efficient and 4 are not Pareto efficient.

2.6 Summary and discussion

In this work an experiment was conducted with the aim to test the performance of the Binary Conditional Contribution Mechanism (BCCM) for public good provision. Since this is the first test a simple binary contribution environment with linear valuations is chosen. In the experiment the BCCM is compared to the standard Voluntary Contribution Mechanism in one setting with complete and one with incomplete information.

In all settings the BCCM leads to significantly higher contribution rates than the VCM. This effect is especially large if only the second half of the experiment is considered. In those periods convergence in many groups of the BCCM is complete and average contribution rates are rather stable at 93% (complete information) or 81% (incomplete information). By comparison, average contribution rates over the same periods under the VCM are 60% (complete information) and 53% (incomplete information). Another important difference between the mechanisms is that in groups playing the BCCM no decline of contributions over time can be observed.

This experiment further gives support for the dynamic model Unexploitable Better Response Dynamics designed in Reischmann (2015a). The model gives an accurate prediction of the long term stable outcomes of the BCCM in the test environment. And all those outcomes are Pareto efficient.

With the apparent experimental success of the BCCM the non-binary Conditional Contribution Mechanism should be tested soon in a follow-up experiment. Thus, next tests should focus on non-binary and/or non-linear environments. Considering the intuitive appeal and simplicity of the message space the Conditional Contribution Mechanisms are further suited to be tested in field experiments.

The BCCM is a new mechanism for public good provision that satisfies individual rationality and incentive compatibility. Furthermore, the mechanism is, compared to many other existing mechanisms, rather simple. With the success of the BCCM in this experimental study it becomes a candidate to finally solve the free-rider problem in a fitting class of public good environments.

Appendix B

This Appendix covers the experiment instructions. They are translations from the German original. The German version can be obtained on request from the author. The different instructions for the four treatments are given in the following order: 1.) VCM, complete information, 2.) CCM, complete information, 3.) VCM, incomplete information 4.) CCM, incomplete information.

Instructions for VCM with complete information

Instructions

Welcome to our experiment! Please read the instructions carefully. Do not talk to your neighbor from now on. Shut down your mobile phone and keep it turned off until the experiment ends. If you have any questions, raise your hands. We will come to you. All participants have got the same instructions.

In the experiment you will be divided in **groups of 5**. The experiment will last for **20 periods**. You will be grouped with the same four players in all periods. The experiment is entirely anonymous. No player will be informed whom he was grouped up with or what payoff any other player obtains.

Points are the currency in the experiment. In every period you start with 10 points. These points will by the end of the period either be added entirely to your private account, or will be invested entirely into a common project. For every player who invests his **10 points** into the project all players obtain **6 points**.

Example 1: You invest you 10 points into the project and 2 other players invested into the project additionally. You will get for your investment and for the investment of the other 2 players 6 points each. Thus you will get $3 \times 6 = 18$ points in total added to your account.

Example 2: You do not invest your 10 points into the project and 2 players invested into the project in total. Your will keep your 10 points and get additionally 6 points each for the investment of the other 2 players. Thus you get $10 + 2 \times 6 = 22$ points in total added to your account.

Every player can choose in every period between two actions:

- You can invest your 10 points into the project.
- Or you can keep your 10 points for yourself.

All players decide simultaneously.

Payoff of all periods

After it was determined who contributes to the project in a given period, all players get the corresponding points added to their account.

Then a new period starts. After 20 periods there will be a questionnaire. After the experiment you will be called to receive your money. You will receive your **earnings for all periods** at a rate of 40 points=1€. The payment will be private and in cash.

Program structure

You obtained a printed example for the structure of the program, which you will use to submit your decision in every period. The screen is divided into three blocks.

The block on the upper left side contains a calculator. Here you can test actions for you and the four other players. Once you select an action for every player the computer will calculate the payoff you would obtain in this case.

In the upper right block you enter the action that will be relevant for your payoff. Below there is a red button. When you push this button you submit your decision and leave the screen. Only when **all** players pushed the button the experiment continues. A clock on the upper right hints at the time in which your decision should be made. **If the time runs out this has no effect.**

From period two on the actions of all players of all previous periods and your payoff in those periods will be displayed in the big block below. In the first period this block will be empty.

The texts in the green frames on the printed example of the program are comments that explain the print. They will not be displayed in the actual program.

Instructions for CCM with complete information

Instructions

Welcome to our experiment! Please read the instructions carefully. Do not talk to your neighbor from now on. Shut down your mobile phone and keep it turned off until the experiment ends. If you have any questions, raise your hands. We will come to you. All participants have got the same instructions.

In the experiment you will be divided in **groups of 5**. The experiment will last for **20 periods**. You will be grouped with the same four players in all periods. The experiment is entirely anonymous. No player will be informed whom he was grouped up with or what payoff any other player obtains.

Points are the currency in the experiment. In every period you start with 10 points. These points will by the end of the period either be added entirely to your private account, or will be invested entirely into a common project. For every player who invests his **10 points** into the project all players obtain **6 points**.

Example 1: You invest you 10 points into the project and 2 other players invested into the project additionally. You will get for your investment and for the investment of the other 2 players 6 points each. Thus you will get $3 \times 6 = 18$ points in total added to your account.

Example 2: You do not invest your 10 points into the project and 2 players invested into the project in total. Your will keep your 10 points and get additionally 6 points each for the investment of the other 2 players. Thus you get $10 + 2 \times 6 = 22$ points in total added to your account.

Every player can **choose** in every period **between six different conditions**:

- 0=Contribute in any case.
- 1=Contribute only if at least one other player contributes, too.
- 2=Contribute only if at least two other players contribute, too.
- 3=Contribute only if at least three other players contribute, too.

- 4=Contribute only if all four other players contribute, too.
- 5=Contribute in no case.

The computer selects the highest amount of players, which can contribute to the project, without violation the condition of any player. These players will then automatically contribute to the project. The other players will not contribute.

Example 1: 3 players choose condition ”1” and the other two players choose condition ”5”. Then those 3 players, who chose condition ”1” will contribute to the project.

Example 2: 3 players choose condition ”3” and the other two players choose condition ”5”. Then no player will contribute to the project.

Payoff of all periods

After it was determined who contributes to the project in a given period, all players get the corresponding points added to their account.

Then a new period starts. After 20 periods there will be a questionnaire. After the experiment you will be called to receive your money. You will receive your **earnings for all periods** at a rate of 40 points=1€. The payment will be private and in cash.

Program structure

You obtained a printed example for the structure of the program, which you will use to submit your decision in every period. The screen is divided into three blocks.

The block on the upper left side contains a calculator. Here you can test conditions for you and the four other players. Once you select a condition for every player the computer will calculate the payoff you would obtain in this case.

In the upper right block you enter the condition that will be relevant for your payoff. Below there is a red button. When you push this button you submit your decision and leave the screen. Only when **all** players pushed the button the experiment continues. A clock on the upper right hints at the time in which your decision should be made. **If the time runs out this has no effect.**

From period two on the conditions of all players of all previous periods and your payoff in those periods will be displayed in the big block below. In the first period this block will be empty.

The texts in the green frames on the printed example of the program are comments that explain the print. They will not be displayed in the actual program.

Instructions for VCM with incomplete information

Instructions

Welcome to our experiment! Please read the instructions carefully. Do not talk to your neighbor from now on. Shut down your mobile phone and keep it turned off until the experiment ends. If you have any questions, raise your hands. We will come to you. All participants have got the same instructions.

In the experiment you will be divided in **groups of 5**. The experiment will last for **20 periods**. You will be grouped with the same four players in all periods. The experiment is entirely anonymous. No player will be informed whom he was grouped up with or what payoff any other player obtains.

Points are the currency in the experiment. In every period you start with 10 points. These points will by the end of the period either be added entirely to your private account, or will be invested entirely into a common project.

At the beginning of the first period every player will be assigned one **type**, which he will keep **for the entire game**.

With a chance of 20% you are type 1 and you do not benefit from the common project. In this case in each period your 10 points will be added to your private account if you do not invest them into the project. And 0 points will be added to your private account if you invest into the project. If any other players invest into the project does not influence your payoff in this case.

With a chance of 80% you are type 2 and you benefit from the common project. In this case in each period your 10 points will be added to your private account as well if you do not invest them into the project, but 6 points will be added to your private account if you invest into the project. Additionally you receive 6 points for every other player, who also invests into the project.

The types are drawn independently, especially different players may thus have different types. Every player gets displayed his type in every period. He does not get to know the types of the other players.

Example 1: You are type 2 and you invest you 10 points into the project and 2 other players invested into the project additionally. You will get for your investment and for the investment of the other 2 players 6 points each. Thus you will get $3 \times 6 = 18$ points in total added to your account.

Example 2: You are type 2 and you do not invest your 10 points into the project and 2 players invested into the project in total. Your will keep your 10 points and get additionally 6 points each for the investment of the other 2 players. Thus you get $10 + 2 \times 6 = 22$ points in total added to your account.

Example 3: You are type 1 and you do not invest your 10 points into the project and 2 players invested into the project in total. Your will keep your 10 points and get no additional points for the investment of the other 2 players. Thus you get 10 points in total added to your account.

Every player can choose in every period between two actions:

- You can invest your 10 points into the project.
- Or you can keep your 10 points for yourself.

All players decide simultaneously.

Payoff of all periods

After it was determined who contributes to the project in a given period, all players get the corresponding points added to their account.

Then a new period starts. After 20 periods there will be a questionnaire. After the experiment you will be called to receive your money. You will receive your **earnings for all periods** at a rate of 40 points=1€. The payment will be private and in cash.

If you are type 1 you will receive 5€ additionally to compensate for your lower earning possibilities.

Program structure

You obtained a printed example for the structure of the program, which you will use to submit your decision in every period. The screen is divided into three blocks.

The block on the upper left side contains a calculator. Here you can test actions for you and the four other players. Once you select an action for every player the computer will calculate the payoff you would obtain in this case.

In the upper right block you enter the action that will be relevant for your payoff. Additionally in this block your type is displayed and whether you benefit from the project. Below there is a red button. When you push this button you submit your decision and leave the screen. Only when **all** players pushed the button the experiment continues. A clock on the upper right hints at the time in which your decision should be made. **If the time runs out this has no effect.**

From period two on the actions of all players of all previous periods and your payoff in those periods will be displayed in the big block below. In the first period this block will be empty.

The texts in the green frames on the printed example of the program are comments that explain the print. They will not be displayed in the actual program.

Instructions for CCM with incomplete information

Instructions

Welcome to our experiment! Please read the instructions carefully. Do not talk to your neighbor from now on. Shut down your mobile phone and keep it turned off until the experiment ends. If you have any questions, raise your hands. We will come to you. All participants have got the same instructions.

In the experiment you will be divided in **groups of 5**. The experiment will last for **20 periods**. You will be grouped with the same four players in all periods. The experiment is entirely anonymous. No player will be informed whom he was grouped up with or what payoff any other player obtains.

Points are the currency in the experiment. In every period you start with 10 points. These points will by the end of the period either be added entirely to your private account, or will be invested entirely into a common project.

At the beginning of the first period every player will be assigned one **type**, which he will keep **for the entire game**.

With a chance of 20% you are type 1 and you do not benefit from the common project. In this case in each period your 10 points will be added to your private account if you do not invest them into the project. And 0 points will be added to your private account if you invest into the project. If any other players invest into the project does not influence your payoff in this case.

With a chance of 80% you are type 2 and you benefit from the common project. In this case in each period your 10 points will be added to your private account as well if you do not invest them into the project, but 6 points will be added to your private account if you invest into the project. Additionally you receive 6 points for every other player, who also invests into the project.

The types are drawn independently, especially different players may thus have different types. Every player gets displayed his type in every period. He does not get to know the types of the other players.

Example 1: You are type 2 and you invest you 10 points into the project and 2 other players invested into the project additionally. You will get for your investment and for the investment of the other 2 players 6 points each. Thus you will get $3 \times 6 = 18$ points in total added to your account.

Example 2: You are type 2 and you do not invest your 10 points into the project and 2 players invested into the project in total. Your will keep your 10 points and get additionally 6 points each for the investment of the other 2 players. Thus you get $10 + 2 \times 6 = 22$ points in total added to your account.

Example 3: You are type 1 and you do not invest your 10 points into the project and 2 players invested into the project in total. Your will keep your 10 points and get no additional points for the investment of the other 2 players. Thus you get 10 points in total added to your account.

Every player can **choose** in every period **between six different conditions**:

- 0=Contribute in any case.
- 1=Contribute only if at least one other player contributes, too.
- 2=Contribute only if at least two other players contribute, too.
- 3=Contribute only if at least three other players contribute, too.
- 4=Contribute only if all four other players contribute, too.
- 5=Contribute in no case.

The computer selects the highest amount of players, which can contribute to the project, without violation the condition of any player. These players will then automatically contribute to the project. The other players will not contribute.

Example 1: 3 players choose condition "1" and the other two players choose condition "5". Then those 3 players, who chose condition "1" will contribute to the project.

Example 2: 3 players choose condition "3" and the other two players choose condition "5". Then no player will contribute to the project.

Payoff of all periods

After it was determined who contributes to the project in a given period, all players get the corresponding points added to their account.

Then a new period starts. After 20 periods there will be a questionnaire. After the experiment you will be called to receive your money. You will receive your **earnings for all periods** at a rate of 40 points=1€. The payment will be private and in cash. If you are type 1 you will receive 5€ additionally to compensate for your lower earning possibilities.

Program structure

You obtained a printed example for the structure of the program, which you will use to submit your decision in every period. The screen is divided into three blocks.

The block on the upper left side contains a calculator. Here you can test conditions for you and the four other players. Once you select a condition for every player the computer will calculate the payoff you would obtain in this case.

In the upper right block you enter the condition that will be relevant for your payoff. Additionally in this block your type is displayed and whether you benefit from the project. Below there is a red button. When you push this button you submit your decision and leave the screen. Only when **all** players pushed the button the experiment continues. A clock on the upper right hints at the time in which your decision should be made. **If the time runs out this has no effect.**

From period two on the conditions of all players of all previous periods and your payoff in those periods will be displayed in the big block below. In the first period this block will be empty.

The texts in the green frames on the printed example of the program are comments that explain the print. They will not be displayed in the actual program.

Appendix C

In addition to instructions subjects had to fill out a slide of comprehension questions. A translation of the German original is given exemplary for the case of the CCM and incomplete information:

Comprehension questions - Experiment PGCCM

You are asked to complete two test questions to check whether you understood the instructions completely.

Choose in the following test question 1 a condition for each player. Choose at least **three different** conditions:

Your condition (player 1): ____

Condition player 2: ____

Condition player 3: ____

Condition player 4: ____

Condition player 5: ____

Underline those players, who would contribute to the project in this case:

Player 1 Player 2 Player 3 Player 4 Player 5

What payoff would you obtain in this period if you are of type 2? ___________

Choose also in the following test question 2 a condition for each player. Choose at least **three different** conditions, such that the number of players, who contribute to the project, differs in test question 1 and 2:

Your condition (player 1): ____

Condition player 2: ____

Condition player 3: ____

Condition player 4: ____

Condition player 5: ____

Underline those players, who would contribute to the project in this case:

Player 1 Player 2 Player 3 Player 4 Player 5

What payoff would you obtain in this period if you are of type 1? ___________

Chapter 3

The non-binary Conditional Contribution Mechanism for public good provision. An Experimental Study *

Abstract

There is still no general solution to the free-rider problem in public good environments. The Conditional Contribution Mechanisms were developed recently as an attempt to solve the problem in repeated public good environments. In a first experiment the Binary Conditional Contribution Mechanism lead to higher contribution rates than the Voluntary Contribution Mechanism, when contribution is binary. In this paper we compare the Conditional Contribution Mechanism to the VCM in a non-binary contribution environment. Additionally we compare both mechanisms to a theoretically flawed but simpler version of the Conditional Contribution Mechanism.

*I thank Jörg Oechssler, Peter Dürsch, Florian Kauffeld, Hannes Koppel, Christoph Brunner, Jonas Hedlund and Christoph Vanberg as well as all participants of the seminars at Heidelberg University for helpful comments.

3.1 Introduction

When public goods are subject to voluntary contribution the incentive structure leads to the well known free-rider problem. A new class of mechanisms, the Conditional Contribution Mechanisms (CCMs), has recently been developed by Reischmann (2015a) as a new attempt to solve this problem in repeated public good environments. The idea of the CCMs is to give agents the possibility to condition their contribution towards the public good on a certain total level of contribution provided by all agents. The mechanism the selects the outcome with the highest total level of contributions that is consistent with all these offers.

In repeated public good environments Better Response Dynamics are a good description of subject behavior (Healy, 2006). Two specific dynamic properties make conditional contribution offers work well under such a better response behavior. First, when contributions to the public good in the current outcome are too low, subjects can choose a cooperative conditional contribution offer without being myopically worse off. In this way inefficient outcomes can be left towards more efficient outcomes under Better Response Dynamics. Second, when contributions in the current outcome have reached an efficient level, subjects can condition their contribution on the current level and thus remove any incentives to free-ride on others' contributions. Therefore, efficient outcomes can be stabilized under Better Response Dynamics. Indeed, under a new variant of Better Response Dynamics introduced in Reischmann (2015a) the CCMs lead only to Pareto efficient outcomes in the long run.

The first experiment with the CCM covered an environment in which contributions are binary (Reischmann, 2015b). In this experiment contribution rates under the binary CCM significantly exceeded contribution rates of the binary Voluntary Contribution Mechanism (VCM) in the later periods. Furthermore, all groups that converged to some outcome under the CCM converged to a Pareto optimal outcome as predicted by the theory.

This paper covers a follow up experiment that tests the CCM in a non-binary contribution environment. In the non-binary version of the CCM all subjects have to offer two (not necessarily different) amounts to contribute to the public good. Both amounts can be coupled with a condition on total contribution. Thus, any contribution offer is only invested into the public good if the corresponding condition can be met. The CCM would be simpler and easier to understand if agents only had to choose one contribution offer and one condition. However, Reischmann (2015a) shows that this mechanism can trap the dynamics on outcomes that are not Pareto optimal. This leads to two questions that this experiment intends to answer. First, can the CCMs significantly increase contributions to public goods compared to the VCM in a non-binary contribution environment, as well? Second, do agents really get trapped in not Pareto optimal outcomes in the simpler version of the mechanism as theory predicts?

Therefore, I compare three mechanisms in a standard linear public goods environment: the standard Voluntary Contribution Mechanism (VCM), the non-binary Conditional Contribution Mechanism (CCM), and the Single Conditional Contribution Mechanism (SCCM) in which only one conditional contribution offer is announced. All mechanisms are played in groups of 5 agents with a homogeneous and linear valuation for the public good. The mechanisms are repeated over 40 periods to give enough time for the dynamic adjustment behavior to converge to an equilibrium. The VCM is chosen as a benchmark since it is the one mechanism that is used most often in real world applications. Thus, finding a mechanism that is simple enough to be applicable in real public good environments and that improves on the performance of the VCM is still one of the most important goals in the research on public goods.

This experiment has two main findings. First the CCMs increase contributions significantly compared to the VCM in periods in which coordination under the CCMs does not fail. Second, coordination fails in almost one in four periods. If those periods are taken into account the CCMs do not lead to a significant increase in contributions. Thus, an improvement in coordination might make the CCMs an option to solve the free-rider problem for repeated non-binary contribution environments. Therefore, the last section of this paper discusses three possible ways to improve coordination without compromising the incentive structure of the CCMs.

3.1.1 Related literature

The related literature can be clustered into three different branches. First, this paper relates to previous work on the CCMs. The CCMs were first introduced in Reischmann (2015a) as a set of mechanisms for repeated public good environments. Reischmann

(2015a) further introduces and motivates a new variant of Better Response Dynamics as solution concept for the theoretical analysis of the CCMs. Under this new solution concept all stable outcomes of the CCMs are Pareto efficient. In the following first experiment on the CCMs Reischmann (2015b) tested the mechanisms in a binary contribution environment once under perfect and once under imperfect information. In both treatments the CCM lead to significantly higher contribution rates in the long run. Furthermore, all long term stable outcomes were in line with the prediction of UBRD.

Second, there are studies on subject behavior in public good environments that are an important motivation for the CCMs. On the one hand there are the experiments from Fischbacher et al. (2001) and Kocher et al. (2008). Both papers show experimentally that there is a significant share of subjects that strictly prefers conditional cooperation over free-riding in public good environments. Nevertheless the good dynamic properties of the theoretical analysis of the CCM hold already under the assumption of selfish agents. Therefore, preferences for conditional cooperation strengthen the argument that the CCM has a rather simple and intuitive message space. Further there is the study of Healy (2006), who finds that Better Response Dynamics explains subject behavior in many public good experiments. This supports the dynamic model used in the theoretical analysis of the CCMs.

Third, there is the branch of the literature that experimentally compares different mechanisms for public good provision. Examples are the studies by Smith (1979), Smith (1980) and Banks et al. (1988), who compare the VCM to the auction mechanism of Smith. Rondeau et al. (1999), Rose et al. (2002) and Rondeau et al. (2005) further test the performance of the VCM against a Provision Point Mechanism. And very recently Morgan and Sefton (2000), Schram and Onderstal (2009), and Corazzini et al. (2010) compare auction and lottery mechanisms to the VCM.

3.1.2 Plan of the paper

The rest of the paper is structured as follows. Section 3.2 introduces the theoretical environment as well as the mechanisms that will be compared in the experiment. In section 3.3 I give a short review of the theoretical predictions of the mechanisms for the given environment. Section 3.4 motivates and presents the experimental design, while section 3.5 presents the results. In section 3.6 I discuss possible ways to improve performance of the tested mechanisms on a theoretical level. Section 3.7 gives a summary and discussion of the paper. Translations of the German instructions and test questions handed out to subjects can be found in Appendix D and E.

3.2 Environment and mechanisms

In this section I give a short introduction to the public good environment that we use in this experiment and the mechanisms that are subject of this experimental study. For an extensive motivation and the thorough theoretical analysis of the Conditional Contribution Mechanisms in more general environments see Reischmann (2015a).

I consider an environment with one private good x and one public good z. In this environment $n = 5$ subjects interact repeatedly and have a constant endowment of $w = 20$ in each period. There is no transfer of money from one period to the next so the entire endowment has to be spent on either the private or the public good in each period. I characterize an outcome in any period in this economy as a vector $z = (z_1, \ldots, z_n)$, where z_i describes the contribution of subject i to the public good. For notational convenience define $\underline{z} = (0, \ldots, 0)$. The payoff functions of all subjects are linearly increasing in the level of the public good.

$$\Pi_i(z) = 20 - z_i + \theta_i \sum_{j=1}^{5} z_j \tag{3.1}$$

Here θ_i is the valuation of agent i for the public good. In the experiment a valuation of $\theta_i = 0.4$ is chosen for all subjects.

There are three treatments that differ in the mechanism which is used to allocate the resources between the public and the private good. In the reference treatment subjects allocate their endowment via the Voluntary Contribution Mechanism (VCM). This means all subjects choose individually and simultaneously how much of their endowment they invest into the public good.

Both other treatments cover versions of the Conditional Contribution Mechanism (CCM). The main treatment covers the CCM as described in Reischmann (2015a). Here all subjects announce two statements of the form "I am willing to contribute an amount α to the public good, if total contributions are at least β". The mechanism checks all those statements and implements of all outcomes z, which satisfy all those conditions, the one with the highest total level of contribution. This can be formalized in the following way. Describe the message of subject i as $m_i = (\alpha_i^1, \beta_i^1, \alpha_i^2, \beta_i^2)$. The set of outcomes which satisfy all conditions can be described by Z^{CCM}:

$$z \in Z^{CCM}(m) \Leftrightarrow \; z_i = 0 \text{ or } \left\{ \exists l_i \in \{1,2\} : \; z_i = \alpha_i^{l_i} \text{ and } \sum_{j=1}^{n} z_j \geq \beta_i^{l_i} \right\}, \; \forall\, i \in I \quad (3.2)$$

Of all those outcomes the outcome function g^{CCM} of the CCM selects the uniquely defined outcome with the highest level of total contributions.

$$g^{CCM}(m) = \text{argmax}_{z \in Z^{CCM}(m)} \sum_{i=1}^{n} z_i. \quad (3.3)$$

This definition of the CCM seems to be unnecessarily complicated. Why announce two conditional contribution statements? Does not one do the trick? One can indeed define the mechanism in which only one such statement is announced and I will test this mechanism in this study as well. I call it here the Single Conditional Contribution Mechanism (SCCM).[1]

From a static perspective the equilibrium outcomes do not change at all when only one message is announced. However, under dynamic considerations the CCM in this form has a substantial advantage over the SCCM from a theoretical perspective as will be shown in the next section. Therefore, there is a trade-off between the CCM and the SCCM. The SCCM has the advantage of simplicity, which may make coordination easier. The CCM grants higher flexibility, which makes deviations that have the aim to improve contributions in the long run less costly. Which effect is more important will be the second focus of this study.

[1] In Reischmann (2015a) this mechanism was called the Natural Extension Mechanism

3.3 Theoretical predictions

This section covers the theoretical predictions of the CCM and SCCM as well as the expectations of the VCM. I discuss Nash equilibrium as a static solution concept and Unexploitable Better Response Dynamics (UBRD) as a dynamic solution concept. All theoretical claims are illustrated by examples. The corresponding theorems, proofs, and the motivation of the solution concept UBRD can be found in (Reischmann, 2015a). Here I just define UBRD before discussing the theoretical predictions of the three mechanisms.

UBRD combines two behavioral assumptions into one model. First subjects are assumed to better respond. I.e. subjects do not switch from their message m_i^t in period t to a message m_i^{t+1} in period $t+1$ if m_i^{t+1} makes agent i worse off if all other subjects repeat what they did in period t.

Second subjects are assumed not to deviate to messages that can be exploited by other agents. This reflects the assumption that each subject takes the possibilities for deviations of other agents into account when choosing his message. And he rejects messages that set incentives for other agents to free-ride on his contributions. In this sense a message m_i^{t+1} is exploitable if there is any message profile m_{-i}^{t+1} such that agent i has to contribute in the outcome of m^{t+1}, but is worse off than in the current outcome under m^t. The formal definition of "exploitable" is given in the following way.

Definition 3.1. Given a message profile m and an outcome $g(m) = z$, a deviation from m_i to m_i' is called *exploitable* if there exists a message profile $m_{-i} \in M_{-i}$ such that $z'(m_{-i}) := g(m_i', m_{-i}) \prec_i z$ and $z_i'(m_{-i}) > 0$. A message m_i' is called *unexploitable*, if it is not *exploitable.*

This results in the following formal model of UBRD.

Definition 3.2. In Unexploitable Better Response Dynamics (UBRD) all agents can adjust their message in every period. Agent i switches in period t to message m_i^t with strictly positive probability if and only if

- m_i^t is a (weak) better response to m^{t-1} and
- m_i^t is unexploitable with respect to $z^{t-1} := g(m^{t-1})$.

The intention of this model is the correct prediction of the long term stable outcomes of the CCMs. The model is not intended to perfectly describe subjects behavior in every period. I will apply this model to the three mechanisms in the next subsections.

3.3.1 Voluntary Contribution Mechanism

Under the assumptions on preferences in this paper free-riding is a strictly dominant strategy for all subjects in the VCM. This implies that there is only one Nash equilibrium. In this Nash equilibrium nobody contributes anything to the public good. Further, free-riding is the only unexploitable better response for any player in any period. Therefore, the dynamic prediction is the same as the static prediction in this case.

However, previous experiments find consistently that contributions in the first period of repeated public good games are about in the middle between the free-rider and the Pareto efficient quantity. Then those contributions decline over time. See Ledyard (1994) for a survey of this literature. Since those findings are persistent this is also the expectation of the results for the VCM in this experiment.

3.3.2 Single Conditional Contribution Mechanism

The Nash equilibrium outcomes of the SCCM are all outcomes that are (weak) Pareto improvements over $\underline{z}$. Such an outcome z can be implemented as Nash equilibrium by letting all subjects offer an individual contribution of z_i under the condition that total contributions are at least $\sum_{j=1}^{n} z_j$. Nobody can benefit individually by deviating from this message profile. This implies that there are multiple efficient as well as inefficient equilibria.

Under UBRD there are many inefficient outcomes of stable states as well. In fact every outcome that is a strict Pareto improvement over $\underline{z}$ is an outcome of a stable state of the dynamics. The following example demonstrates this.

Example 3.1. *Consider the experimental environment of* $n = 5$ *subjects with an endowment of* $w = 20$*. Assume that in the current period all subjects send the message* $m_i = (10, 50)$*. Therefore, all subjects offer to contribute* 10 *if total contributions are at least* 50*. This leads to the outcome* $z = (10, 10, 10, 10, 10)$*, which is a strict Pareto* *improvement over* $\underline{z}$. *Assume now that subject* 1 *wants to deviate. If subject 1 reduces his contribution offer to anything lower than* 10 *the conditions of subjects 2 to 5 will be violated and they will no longer contribute anything to the public good. Therefore, this is not a better response for subject 1. If subject 1 increases his contribution offer while keeping his condition fulfilled he increases his own contribution without altering the contribution of anyone else. This makes him worse off as well. If subject 1 increases his contribution combined with a condition that can not be satisfied, the outcome will be* $\underline{z}$. *All remaining messages offer a contribution of* 10. *However, if this is combined with any condition larger than* 50 *the outcome is again* $\underline{z}$. *And if it is combined with a condition smaller than* 50 *the message is exploitable. Therefore* $m_i = (10, 50)$ *is the only unexploitable better response. Thus, this is one example of steady state of UBRD, in which the outcome is not Pareto optimal.*

In the same way as in the example subjects can get trapped in a steady state with any other outcome that is a strict Pareto improvement over $\underline{z}$. Of course if subjects realize the potential long term gains it might be reasonable for them to break such steady states and accept the short term loss. This would lead to only efficient steady states and then the SCCM should be preferred over the CCM for its simplicity. However, whether this would actually happen is an empirical question. Therefore, the SCCM is included as a treatment in this study to answer this question.

3.3.3 Conditional Contribution Mechanism

The CCM has exactly the same Nash equilibrium outcomes as the SCCM. Therefore the two mechanisms do not differ at all with respect to static equilibrium properties. However, with respect to dynamic equilibrium properties there is a significant difference.

Under UBRD all outcomes of recurrent classes of the CCM are Pareto efficient and strict Pareto improvements over $\underline{z}$. The reason for this result is that in the CCM any outcome that is not Pareto efficient can be left towards Pareto improvements in line with UBRD. This can be illustrated by using a similar example as above.

Example 3.2. *Assume that in the experimental environment all* 5 *subjects currently offer to contribute* 10 *if and only if total contributions are at least* 50. *This corresponds to messages* $m_i = (10, 50, 10, 50)$ *for all agents. Now assume agent 1 deviates to the* *message* $m'_1 = (10, 50, 20, 100)$. *This deviation does not change the outcome, therefore it is a better response. Further if the second condition should actually be satisfied total contributions would increase from* 50 *to* 100 *making subject 1 better off even though he has to contribute more. Thus,* m'_i *is unexploitable. However, if all subjects switch to* $m'_i = (10, 50, 20, 100)$ *the outcome will switch from* $z = (10, 10, 10, 10, 10)$ *to* $z' = (20, 20, 20, 20, 20)$ *and all subjects will be strictly better off.*

Therefore, the theoretical prediction for the CCMs is that the agents will reach some stable state eventually. In case of the SCCM it will depend on the ability of subjects to evaluate the trade off between long term and short term gains if those outcomes will be Pareto efficient. In case of the CCM those outcomes should always be Pareto efficient. In previous experiments with the binary version of the CCM the model of UBRD already predicted the long term stable outcomes correctly (Reischmann, 2015b).

3.4 Experimental design

These experiments were designed to answer the following two questions. Do the Conditional Contribution Mechanisms lead to more efficient contributions than the VCM in the selected environment? Are the better theoretical dynamic properties of the CCM more important than the simplicity of the SCCM?

In the experiment 110 subjects were randomly grouped in 22 groups of 5 people. All groups stayed together for the entire experiment. Each group played a public good game with one of the three mechanisms. Eight groups played the SCCM, seven groups played the VCM, and seven groups played the CCM. The stage game mechanism was repeated over 40 periods. 40 periods were chosen since already in the experiments on the BCCM about 10 periods were needed to achieve coordination on stable outcomes. Since the CCM is significantly more complicated it is reasonable to assume that convergence will take longer. 40 periods should be long enough to observe whether convergence does take place.

All experiments were conducted at the experimental lab of the Alfred-Weber Institute of Heidelberg University. At the beginning of the experiments subjects were randomly placed inside the lab and handed out instructions and test questions. The experiments were conducted in German. English translations of the instructions and test questions can be found in Appendix D and E. In the test questions subjects are asked to choose messages for the corresponding mechanism for 5 subjects and then calculate the mechanism outcome by hand. In the case of the CCM a hint for a fast way to calculate the outcome was given on the sheet. Once all subjects solved the test question correctly and there were no more questions the experiment started.

The experiment was programmed in the software z-Tree (Fischbacher, 2007). In each period subjects were asked to enter their message. They further had access to a calculator in which they could enter messages for all subjects and let the calculator determine their contribution as well as their payoff in this case. From period two onwards subjects were displayed a history that contains the messages of all subjects of previous rounds as well as their own payoff in this period.

After the 40 periods were finished subjects were asked to complete a short questionnaire asking for personal characteristics such as prior knowledge of game theory. Once all subjects finished the questionnaire they were paid their earnings in private and in cash. The experiments lasted between 60 and 90 minutes and subjects earned on average 12.33€.

3.5 Experimental results

The analysis of the experimental data focuses on two different aspects. First, whether or not average contribution rates in groups are significantly different in the three mechanism. And second, whether or not the dynamic predictions of UBRD can be observed in the data of the experiments on the CCMs. In particular, do the bad dynamic properties of the SCCM lead to suboptimal stable outcomes as predicted by the model UBRD? And are all long term stable outcomes under the CCM Pareto efficient?

3.5.1 Contribution rates

I want to start with the simple observation that the experimental data with respect to the VCM is in line with previous findings in the experimental literature. Average contributions in period 1 are 68% and contributions decline over time. The decline is estimated by a Matched-pairs Wilcoxon-Ranksum Test comparing average contributions over periods 1 to 20 with those from period 21 to 40 for each group. This leads to a p-value of 0.018.

The most important question that this study answers is concerning the comparison of contribution rates between the three mechanisms. Therefore, I look at average contribution rates over either all periods, the last 20 periods, or the last 10 periods in all groups. The second and third analysis has the intention of comparing contribution rates once a certain level of dynamic adjustments took place. I compare these averages between each set of two mechanisms with a Wilcoxon-Ranksum Test. Contribution rates for the CCM and for the SCCM are lower in the beginning, but higher in the last 20 and 10 periods than in the VCM as can be seen in figure 3.1. However, it turns out that non of those differences are statistically significant (all p-values higher than 0.1). This implies that there is no statistically significant difference in actual contributions.

Result 3.3. There are no significant differences in average contribution rates between the three mechanisms.

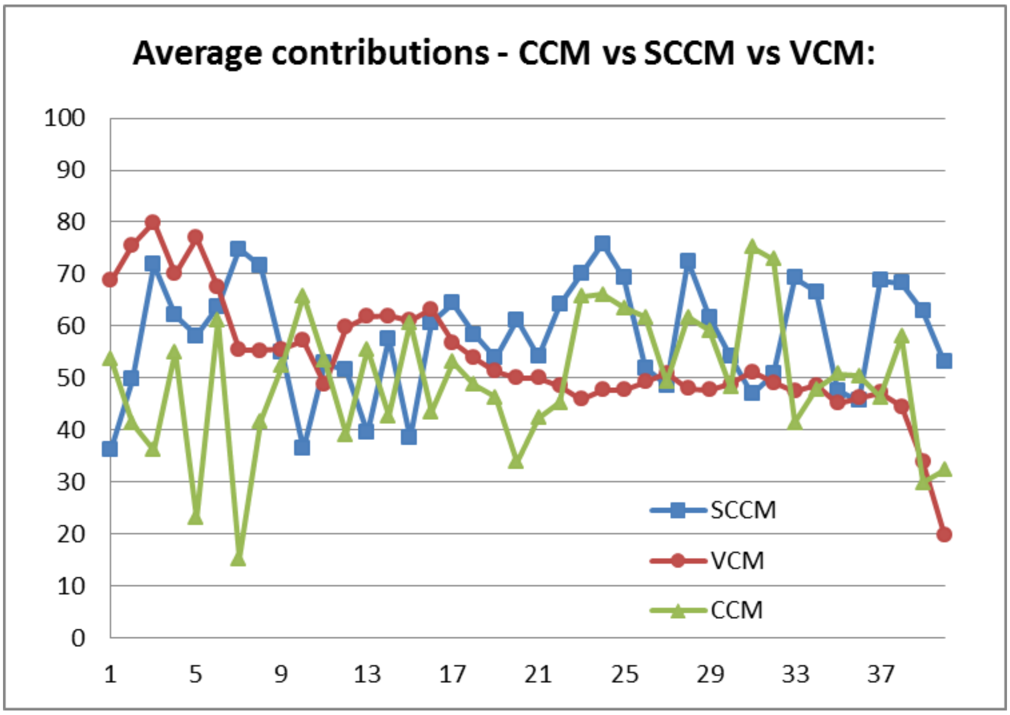

FIGURE 3.1: Comparison of average contributions over all groups

The main reason for these insignificant results is that, while the contribution offers are higher, the CCMs suffer from a lack of coordination in many periods. In fact in about 24% of periods actual contribution is zero although contribution offers are similar to all other periods, because no coordination of offers and conditions was accheived. This hints to the fact that contribution rates in the CCMs might be significantly higher than in the VCM in those periods in which contribution does not fail. This is indeed true for the SCCM in the later periods. I compare the contribution rates between the SCCM and the VCM and evaluate for the SCCM only periods in which total contribution is bigger than zero. Then average contributions are significantly higher in the SCCM than the VCM in the last 20 (p=0.08) and the last 10 periods (p=0.08). However, the same analysis comparing the CCM with the VCM does not lead to significant results. A closer look at individual groups playing the CCM shows that subjects don't manage to improve as a group. The reason for this might be that choosing a strategically good better response becomes too difficult when faced with a total of eight conditional contribution offers of the other four players.

Result 3.4. In periods in which coordination does not fail entirely only the SCCM leads to significantly higher contribution rates than the VCM in the last 20 and the last 10 periods.

Of course this comparison is somewhat unfair and in no way an argument to use the SCCM in its current form. However, these results show that the SCCM leads agents to higher offers in general. And if we can find a way to improve coordination without altering the incentive structure of the SCCM there is a good chance that this mechanism can increase efficiency in repeated public good environments significantly.

3.5.2 Unexploitable Better Response Dynamics

The second question of this paper concerns the dynamic solution model Unexploitable Better Response Dynamics. This model was introduced in the theoretical analysis of the CCMs (Reischmann, 2015a) and predicted the long term stable outcomes of the binary CCM correctly in a previous experimental study (Reischmann, 2015b). For this experiment UBRD makes the following predictions for the two mechanisms. All long term stable outcomes under the CCM are Pareto optimal. All long term stable outcomes in the SCCM are strict Pareto improvements over $\underline{z}$.

Convergence must be defined in some way in the data to answer those questions experimentally. I say that a group "converged to a stable outcome" if at least for of the last five outcomes are identical. With this definition five out of the 15 groups that use one of the two versions of the CCM converged to some outcome, three groups for the SCCM and two groups four the CCM. In the three groups with the SCCM as well as one group with the CCM this outcome is the most efficient Pareto optimal outcome. And in one remaining group with the CCM this outcome is $\underline{z}$. In the last group contribution offers are quite high, but in the last four periods coordination fails entirely and actual contributions are zero in those four periods.

These observations show two results. First, in about one in four groups subjects manage to coordinate for multiple periods on a Pareto optimal outcome. This indicates the UBRD predicts correctly that those outcomes are stable states of the mechanisms.

Result 3.5. Pareto efficient outcomes are stable states of the CCMs as predicted by the model UBRD.

Second, groups do not get stuck on any inefficient outcomes that are stable states of UBRD. However, in at least one group we observe that subjects deliberately choose to break coordination in order to coordinate on a more efficient long term outcome. This accounts for some of the periods in which coordination fails in the SCCM.

Result 3.6. There is no support for the prediction of UBRD that the SCCM has also inefficient stable outcomes.

3.5.3 Summary

Of the two CCMs tested in this experimental study the SCCM seems to be more promising. It leads to significantly higher contribution rates than the VCM in those periods in which coordination does not fail. Thus, finding ways to improve coordination in the SCCM can be one next step in the research on CCMs. I present some ideas on this matter in the next section.

3.6 Three ways to improve coordination

The experimental results demonstrate that the main issue of the SCCM is a coordination problem. Therefore, I conclude this paper by proposing three different possibilities to facilitate coordination.

3.6.1 Possibility I: Use communication

The first possibility to improve coordination is the introduction of communication. This is certainly the most simple way to improve coordination, whenever it is applicable. Already in experiments with the VCM subjects have been shown to achieve more efficient contribution rates when communication is possible (Isaac et al., 1985). Since communication in such an environment does not change the incentive structure of the mechanisms this increase in contributions can not be related to coordination in the VCM. However, communication in form of promises can be more than cheap talk. As Vanberg (2008) suggests subjects may have direct preferences for promise keeping.

However, in the SCCM it does not even matter if subjects have preferences for keeping their promise. Once subjects manage to coordinate on a certain outcome by means of communication nobody has an incentive to deviate from that agreement. He would just violate the conditions of everyone else leading to a break down of cooperation. Therefore, we can expect communication to have an even more positive effect on the SCCM than on the VCM.

However, there are two reasons that make communication less appealing. The first reason is that communication in real applications is often costly. Sometimes with regards to monetary costs, but almost always with regards to costs in terms of time and effort. The second reason is that communication is just not always practicable. Especially when groups get bigger communication between all participants might just be impossible. Nevertheless, for small groups with low communication costs communication is certainly an important option to facilitate coordination.

3.6.2 Possibility II: Use an extensive-form

The second possibility to improve coordination is to change the stage game to an extensive form game without changing the message spaces for any player. This has the advantage that the players moving later in the game already know the conditions of all previous players. However, there is the disadvantage that the game is no longer symmetric and therefore the first mover has a significant advantage. However, when subgame perfect equilibrium is considered as the natural solution concept the outcome will still be Pareto efficient. The structure of the equilibrium can easily be observed in an example.

Example 3.3. *Consider the experimental environment with* $n = 5$ *subjects, endowments of* $w = 20$*, and valuations for the public good of* $\theta_i = 0.4$*. For simplicity call the subject that moves first subject 1 and the subject that moves last subject 5. Further assume that whenever subjects are indifferent they choose the option that leads to higher contribution levels.*[2] *Then apply backward induction. Subject 5 knows all contribution offers and conditions of all other subjects. Thus, he will simply choose of all contribution levels the one that gives him the highest payoff. Importantly, if nobody chose to contribute anything so far he will contribute zero as well. Therefore, he basically can choose between satisfying the conditions of subjects 1 to 4 in the cheapest way or the outcome* $\underline{z}$*. Subject 4 faces a similar problem. Since the combined benefit of subjects 4 and 5 is only* $\theta_4 + \theta_5 = 0.8$*, there is no Pareto improvement over* z_0 *if only subjects 4 and 5 contribute. Therefore also subject 4 faces the trade off to either satisfy the conditions of all previous players, anticipating the optimal contribution decision of subject 5, or the outcome* $\underline{z}$*. Subject 3 finally faces a different choice. If he chooses not to satisfy the conditions of subject 1 and 2 he can choose the message* $m_3 = (10, 50)$*. This makes subjects 4 and 5 indifferent between choosing* $m_{4/5} = (20, 50)$ *and* $m_{4/5} = (0, 0)$*. And we assumed that they would choose the first option. Therefore, subject 3 can decide between the optimal way of satisfying the conditions of 1 and 2 and the outcome* $z = (0, 0, 10, 20, 20)$ *giving him a payoff of* 30*. Thus subject 3 will only satisfy conditions of players 1 and 2 if this leads to a payoff of at least 30. Subject 2 has no way of incentivizing player 3 on his own to increase contributions even more, because subjects 4 and 5 already contribute the maximum and valuations of subject 2 and 3 only add up to* 0.8 *again. This leaves the decision of subject 1. The best he can offer is* $m_1 = (5, 75)$*. This makes subject 2 indifferent between* $m_2 = (10, 75)$ *and* $m_2 = (0, 0)$ *where he chooses the first one. This makes subject 3 indifferent between* $m_3 = (20, 75)$ *and* $m_3 = (10, 50)$ *again choosing the first one. Thus, on the equilibrium path in the subgame perfect equilibrium subjects choose* $m^* = (5, 75; 10, 75; 20, 75; 20, 75; 20, 75)$*, with outcome* $z^* = (5, 10, 20, 20, 20)$*.*

The example demonstrates the Pareto efficiency of the outcome as well as the first mover

[2] This assumption is not critical. Assuming that subjects instead choose the outcome with lower contributions in case of indifference only reduces the first mover advantage to some extent.

advantage in this structure. The mechanism can of course be made fair ex ante by having a fair random draw over the order of players. However, this does not make this mechanism fair ex post, which would be desirable. Further, as in the case of communication an extensive form game becomes more difficult to implement the more players are participating in the mechanism. One way to go here might be to have an extensive form game without a fixed order. E.g. such that players can commit to their messages on an online platform and all players can inform themselves on all messages committed so far at any given time.

3.6.3 Possibility III: Keep gains of previous coordination

The third possibility to improve coordination is to change the outcome generation of the mechanism to a two stage process. The general idea is, that whenever coordination fails the mechanism proposes a "better" outcome which can get implemented if everyone prefers it. Basically the mechanism coordinates your messages for you. Thus the mechanism would calculate the standard outcome of the CCMs and then let agents vote between this outcome and the proposed "better" outcome. Only if all agents vote for the "better" outcome it gets implemented. Otherwise the outcome will be the standard outcome as described in this paper.

However, there is one big problem with this approach. The mechanism needs to know somehow what is a "better" outcome without knowing subjects preferences. Therefore, the only information that can be used in the stage game is the message profile sent by agents. If the proposed alternate outcome depends on agents' messages this sets incentives to break coordination on purpose in a way that the most beneficial alternate outcome gets proposed. Since breaking the coordination usually leads to $\underline{z}$ agents would vote between $\underline{z}$ and some coordinated outcome that the deviator can influence. Therefore, I can see no way to define a "better" outcome depending on the message profile that does not ruin the incentive structure of the mechanism.

Thus, this proposed alternate outcome cannot depend on the message profile, either. In a one-shot game there is nothing left that can be used. However, in the repeated settings that are considered to be suited for the CCMs there is the history of past play. Thus, I propose to use the outcome of the last period as an alternate outcome. This makes it possible, even in the simple SCCM, to choose your messages in an attempt to realize further Pareto improvements. And if this leads to a failure of coordination the gains from previous periods can be saved by voting for this outcome. Since this alternate outcome does not depend on any current message this change does not alter the static incentive structure of the mechanism. And it even strengthens the argument for unexploitable messages in the dynamic model.

3.7 Summary and discussion

The focus of this study is the experimental test of two versions of the Conditional Contribution Mechanism (CCM) for public good provision. In one version, the Single Conditional Contribution Mechanism (SCCM), agents announce one contribution offer and one condition. In the second version agents announce two such pairs. Both mechanisms are further compared to the performance of the Voluntary Contribution Mechanism, which is still used in most applications for its simplicity.

The experimental study answers two questions. First, do the CCMs provide a more efficient level of the public good than the VCM? The straight forward answer to this question is no. Even if only the last 20 or the last 10 periods are taken into consideration the difference in contribution rates is not significant. This implies that in the tested environment the CCMs in the tested configuration do not improve the level of contributions to public goods. The main reason for this failure is that in both CCMs agents' conditional contribution offers do not match well leading to a contribution of zero in about 24% of all periods.

However, if we compare only contribution rates in periods in which coordination does not fail entirely contributions in the SCCM are significantly higher than in the VCM. This observation does not make the result of the SCCM any better, but it demonstrates that there is a potential of significant improvements if the problem of coordination of conditional contribution offers can be solved. Therefore, the last section of this study offers three possible ways to improve coordination in the SCCM as a hint for future research. In one shot environments coordination can be improved by adding communication, or changing the stage game to an extensive form game. If the environment is a repeated public good setting I propose to vote after every period whether the outcome of the current message profile shall be implemented or whether the outcome of last period shall be used. The outcome of the last period is only used if all players favor this option.

The second question that this study answers is whether or not the predictions of UBRD can be observed in the data of the experiments on the CCMs. For the SCCM UBRD predicts that if the mechanism reaches a stable outcome which is not Pareto efficient, then this outcome will never be left. In the experiment this kind of behavior was not observed. Either subjects did not coordinate on such an outcome at all or they deliberately chose to leave it. This lead to short term losses but long term gains. Further, more than one in four groups converge to the socially most desirable outcome. This supports the prediction of UBRD that Pareto efficient outcomes are stable in the long run.

Summing up, the Conditional Contribution Mechanisms did not directly outperform the VCM in terms of contributions in this non-binary environment. The main problem of the mechanisms is a failure in coordination that occurs in about one out of every four periods. The main goal for future research on the CCMs must therefore be the search for ways to improve coordination in a reliable way.

Appendix D

This Appendix covers the instructions of the experiments. They are translations from the German original. The German version can be obtained on request from the author. The different instructions for the three treatments are given in the following order: 1.) Voluntary Contribution Mechanism 2.) Single Conditional Contribution Mechanism 3.) Conditional Contribution Mechanism.

Instructions for the VCM

Instructions

Welcome to our experiment! Please read the instructions carefully. Do not talk to your neighbor from now on. Shut down your mobile phone and keep it turned off until the experiment ends. If you have any questions, raise your hands. We will come to you. All participants have got the same instructions.

In the experiment you will be divided in **groups of 5**. The experiment will last for **40 periods**. You will be grouped with the same four players in all periods. The experiment is entirely anonymous. No player will be informed whom he was grouped up with or what payoff any other player obtains.

Points are the currency in the experiment. In every period you start with 20 points. These points will by the end of the period either be added directly to your private account, or will be invested into a common project. For every point that is invested into the project by any player all players obtain **0.4 points**.

Example 1: You invest 15 points into the project and additionally 2 other players invested 15 points each into the project. You will get for your investment and for the investment of the other 2 players $45 \times 0.4 = 18$ points. Additionally you keep 5 of your 20 points. Thus, you will get $18 + 5 = 23$ points in total added to your account.

Every player can choose in every period how many points he wants to invest into the project. He can choose any natural number between 0 and 20.

All players decide simultaneously.

Payoff of all periods

After it was determined who contributes how much to the project in a given period, all players get the corresponding points added to their account.

Then a new period starts. After 40 periods there will be a questionnaire. After the experiment you will be called to receive your money. You will receive your **earnings for all periods**. For every point you will receive 0.01€. The payment will be private and in cash.

Program structure

You obtained a printed example for the structure of the program, which you will use to submit your decision in every period. The screen is divided into three blocks.

The block on the upper left side contains a calculator. Here you can test actions for you and the four other players. Once you selected an action for every player you can press the button "Calculate payoff!" and the computer will calculate the payoff you would obtain in this case.

In the upper right block you enter the action that will be relevant for your payoff. Below there is a red button. When you push this button you submit your decision and leave the screen. Only when **all** players pushed the button the experiment continues. A clock on the upper right hints at the time in which your decision should be made. **If the time runs out this has <u>no</u> effect.**

From period two on the actions of all players of all previous periods and your payoff in those periods will be displayed in the big block below. In the first period this block will be empty.

The texts in the green frames on the printed example of the program are comments that explain the print. They will not be displayed in the actual program.

Instructions for the SCCM

Instructions

Welcome to our experiment! Please read the instructions carefully. Do not talk to your neighbor from now on. Shut down your mobile phone and keep it turned off until the experiment ends. If you have any questions, raise your hands. We will come to you. All participants have got the same instructions.

In the experiment you will be divided in **groups of 5**. The experiment will last for **40 periods**. You will be grouped with the same four players in all periods. The experiment is entirely anonymous. No player will be informed whom he was grouped up with or what payoff any other player obtains.

Points are the currency in the experiment. In every period you start with 20 points. These points will by the end of the period either be added directly to your private account, or will be invested into a common project. For every point that is invested into the project by any player all players obtain **0.4 points**.

Example 1: You invest you 10 points into the project and additionally 2 other players invested 5 points each into the project. You will get for your investment and for the investment of the other 2 players $20 \times 0.4 = 8$ points. The 10 points that you did not invest are added directly to your account. Thus you get $8 + 10 = 18$ points in total.

Example 2: You invest you 15 points into the project and additionally 2 other players invested 15 points each into the project. You will get for your investment and for the investment of the other 2 players $45 \times 0.4 = 18$ points. The 5 points that you did not invest are added directly to your account. Thus you get $18 + 5 = 23$ points in total.

Every player can propose in every period a whole-number amount between 0 and 20 points that shall be invested into the project. Further there can be set a condition for this contribution. This condition sets a total amount that has to be invested by all players together (yourself included) into the project such that your contribution is invested as well.

After all players simultaneously choose an amount and a condition, the computer calculates the highest amount possible, which can be invested into the project. Each player will invest his proposed amount if and only if his condition makes it possible.

Example 1: You chose 5 points as potential amount. And 20 points as condition. This is like saying: I am willing to invest 5 points into the project if total contributions to the project are at least 20 points. If the other players chose amounts and conditions such that the computer calculates a maximal possible contribution to the project of less than 20 points you will not contribute to the project at all. If the computer calculates an amount between 20 and 100 points you invest 5 points into the project.

Example 2: You chose 0 points as potential amount. And 0 points as condition. This is like saying: I am willing to invest 0 points into the project if total contributions to the project are at least 0 points. No matter what the other players choose in this case you will not contribute to the project at all and keep your 20 points in your account.

Example 3: You chose 10 points as potential amount. And 10 points as condition. This is like saying: I am willing to invest 10 points into the project if total contributions to the project are at least 10 points. Since your proposed 10 points already satisfy your own condition you will contribute in any case 10 points to the project, no matter how much the other players invest.

Payoff of all periods

After it was determined who contributes how much to the project in a given period, all players get the corresponding points added to their account.

Then a new period starts. After 40 periods there will be a questionnaire. After the experiment you will be called to receive your money. You will receive your **earnings for all periods** at an exchange rate of 100 points = 1. The payment will be private and in cash.

Program structure

You obtained a printed example for the structure of the program, which you will use to submit your decision in every period. The screen is divided into three blocks.

The block on the upper left side contains a calculator. Here you can test actions for you and the four other players. Once you selected an action for every player you can press the button "Calculate payoff!" and the computer will calculate the payoff you would obtain in this case as well as the amount you will invest.

In the upper right block you enter your amount and your condition that will be relevant for your payoff. Below there is a red button. When you push this button you submit your decision and leave the screen. Only when **all** players pushed the button the experiment continues. A clock on the upper right hints at the time in which your decision should be made. **If the time runs out this has <u>no</u> effect.**

From period two on the amounts and conditions of all players of all previous periods and your payoff in those periods will be displayed in the big block below. In the first period this block will be empty.

The texts in the green frames on the printed example of the program are comments that explain the print. They will not be displayed in the actual program.

Instructions for the CCM

Instructions

Welcome to our experiment! Please read the instructions carefully. Do not talk to your neighbor from now on. Shut down your mobile phone and keep it turned off until the experiment ends. If you have any questions, raise your hands. We will come to you. All participants have got the same instructions.

In the experiment you will be divided in **groups of 5**. The experiment will last for **40 periods**. You will be grouped with the same four players in all periods. The experiment is entirely anonymous. No player will be informed whom he was grouped up with or what payoff any other player obtains.

Points are the currency in the experiment. In every period you start with 20 points. These points will by the end of the period either be added directly to your private account, or will be invested into a common project. For every point that is invested into the project by any player all players obtain **0.4 points**.

Example 1: You invest you 10 points into the project and additionally 2 other players invested 5 points each into the project. You will get for your investment and for the investment of the other 2 players $20 \times 0.4 = 8$ points. The 10 points that you did not invest are added directly to your account. Thus you get $8 + 10 = 18$ points in total.

Example 2: You invest you 15 points into the project and additionally 2 other players invested 15 points each into the project. You will get for your investment and for the investment of the other 2 players $45 \times 0.4 = 18$ points. The 5 points that you did not invest are added directly to your account. Thus you get $18 + 5 = 23$ points in total.

Every player can propose in every period two whole-number amounts between 0 and 20 points that shall be invested into the project. Further there can be set a condition for each amount. This condition sets a total amount that has to be invested by all players together (yourself included) into the project such that each of your amounts is invested as well.

After all players simultaneously choose two amounts and conditions, the computer calculates the highest amount possible, which can be invested into the project. Each player will invest at most one of his proposed amounts into the project. If his condition makes it possible he invests the higher amount.

Example 1: You chose 5 and 15 points as potential amounts. And 20 and 60 points as conditions. This is like saying: I am willing to invest 5 points into the project if total contributions to the project are at least 20 points. If total contributions are at least 60 points I an willing to invest 15 points into the project. If the other players chose amounts and conditions such that the computer calculates a maximal possible contribution to the project of less than 20 points you will not contribute to the project at all. If the computer calculates an amount between 20 and 59 points you invest 5 points into the project. And if the computer calculates an amount of 60 or more points you will invest 15 points into the project.

Example 2: You chose 10 and 20 points as potential amount. And 10 and 100 points as conditions. This is like saying: I am willing to invest 20 points into the project if total contributions to the project are at least 100 points. If total contributions are less than 100 points I will invest 10 points in any case into the project. If the other players chose amounts and conditions such that the computer calculates a total amount of less than 100 points you will invest 10 points into the project. If the computer calculates an amount of 100 points you will invest 20 points into the project.

Example 3: You chose 0 and 0 points as potential amounts. And 0 and 0 points as conditions. This is like saying: I am not willing at all to invest into the project. No matter what amounts and conditions the other players choose in this case you will not contribute to the project at all and keep your points in your account.

Payoff of all periods

After it was determined who contributes how much to the project in a given period, all players get the corresponding points added to their account.

Then a new period starts. After 40 periods there will be a questionnaire. After the experiment you will be called to receive your money. You will receive your **earnings for all periods** at an exchange rate of 100 points = 1. The payment will be private and in cash.

Program structure

You obtained a printed example for the structure of the program, which you will use to submit your decision in every period. The screen is divided into three blocks.

The block on the upper left side contains a calculator. Here you can test amounts and conditions for you and the four other players. Once you selected amounts and conditions for every player you can press the button "Calculate payoff!" and the computer will calculate the payoff you would obtain in this case as well as the amount you will invest.

In the upper right block you enter your amounts and your conditions that will be relevant for your payoff. Below there is a red button. When you push this button you submit your decision and leave the screen. Only when **all** players pushed the button the experiment continues. A clock on the upper right hints at the time in which your decision should be made. **If the time runs out this has no effect.**

From period two on the amounts and conditions of all players of all previous periods and your payoff in those periods will be displayed in the big block below. In the first period this block will be empty.

The texts in the green frames on the printed example of the program are comments that explain the print. They will not be displayed in the actual program.

Appendix E

In addition to instructions subjects had to fill out a slide of comprehension questions. A translation of the German original is given exemplary for the case of the SCCM:

Comprehension question - Experiment PGCCM

You are asked to complete a test questions to check whether you understood the instructions completely.

Choose in the following test question one amount and one condition for each player. Choose at least **three different** amounts and conditions:

Your amount (player 1): ____

Your condition (player 1): ____

Amount player 2: ____

Condition player 2: ____

Amount player 3: ____

Condition player 3: ____

Amount player 4: ____

Condition player 4: ____

Amount player 5: ____

Condition player 5: ____

Underline those players, who would contribute to the project in this case:

Player 1 Player 2 Player 3 Player 4 Player 5

What are the total contributions that will be invested into the project in this case?

____

What payoff would you obtain in this case in this period?

____

Bibliography

Banks, J. S., Plott, C. R., and Porter, D. P. (1988). An experimental analysis of unanimity in public goods provision mechanisms. *The Review of Economic Studies*, 55(2):pp. 301–322.

Burger, N. E. and Kolstad, C. D. (2009). Voluntary public goods provision, coalition formation, and uncertainty. Working Paper 15543, National Bureau of Economic Research.

Cabrales, A. and Serrano, R. (2011). Implementation in adaptive better-response dynamics: Towards a general theory of bounded rationality in mechanisms. *Games and Economic Behavior*, 73(2):360 – 374.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, 14(1):47–83.

Chen, Y. (2008). Incentive-compatible mechanisms for pure public goods: A survey of experimental research. In Handbook of Experimental Economics Results, Vol. 1, ed. Charles R. Plott and Vernon L. Smith, pages 625 – 643. Elsevier.

Clarke, E. (1971). Multipart pricing of public goods. *Public Choice*, 11:17–33.

Corazzini, L., Faravelli, M., and Stanca, L. (2010). A prize to give for: An experiment on public good funding mechanisms*. *The Economic Journal*, 120(547):944–967.

Falkinger, J., Fehr, E., Gächter, S., and Winter-Ebmer, R. (2000). A simple mechanism for the efficient provision of public goods: Experimental evidence. *The American Economic Review*, 90(1):247–264.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):pp. 817–868.

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

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, 71(3):397 – 404.

Groves, T. and Ledyard, J. (1977). Optimal allocation of public goods: A solution to the "free rider" problem. *Econometrica*, 45(4):783–809.

Healy, P. J. (2006). Learning dynamics for mechanism design: An experimental comparison of public goods mechanisms. *Journal of Economic Theory*, 129(1):114 – 149.

Isaac, R. M., McCue, K. F., and Plott, C. R. (1985). Public goods provision in an experimental environment. *Journal of Public Economics*, 26(1):51 – 74.

Jackson, M. and Moulin, H. (1992). Implementing a public project and distributing its cost. *Journal of Economic Theory*, 57(1):125 – 140.

Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J., and Sutter, M. (2008). Conditional cooperation on three continents. *Economics Letters*, 101(3):175 – 178.

Ledyard, J. O. (1994). Public goods: A survey of experimental research. Public Economics 9405003, EconWPA.

Lindahl, E. (1919). Just taxation - A positive solution (E. Henderson, trans.). In R.A. Musgrave, A. P., editor, *Classics in the theory of public finance*, pages 98–123. Macmillan, London. 1958.

Morgan, J. and Sefton, M. (2000). Funding public goods with lotteries: Experimental evidence. *The Review of Economic Studies*, 67(4):785–810.

Owen, G. (1982). *Game Theory*, chapter XIII Games without side payments, pages 288 – 328. Academic Press.

Reischmann, A. (2015a). Conditional Contribution Mechanisms for the provision of public goods. *Heidelberg Discussionpaper*, (586). http://www.uni-heidelberg.de/fakultaeten/wiso/awi/forschung/discussionpaper.html.

Reischmann, A. (2015b). Voluntary vs. Conditional Contribution Mechanism - An experimental study. *Heidelberg Discussionpaper*, (587). http://www.uni-heidelberg.de/fakultaeten/wiso/awi/forschung/discussionpaper.html.

Rondeau, D., Poe, G. L., and Schulze, W. D. (2005). VCM or PPM? a comparison of the performance of two voluntary public goods mechanisms. *Journal of Public Economics*, 89(8):1581 – 1592.

Rondeau, D., Schulze, W. D., and Poe, G. L. (1999). Voluntary revelation of the demand for public goods using a provision point mechanism. *Journal of Public Economics*, 72(3):455 – 470.

Rose, S. K., Clark, J., Poe, G. L., Rondeau, D., and Schulze, W. D. (2002). The private provision of public goods: tests of a provision point mechanism for funding green power programs. *Resource and Energy Economics*, 24:131 – 155. Economist and editor George Tolley - A special issue in his honour.

Schram, A. J. and Onderstal, S. (2009). Bidding to give: An experimental comparison of auctions for charity*. *International Economic Review*, 50(2):431–457.

Smith, V. L. (1979). An experimental comparison of three public good decision mechanisms. *The Scandinavian Journal of Economics*, 81(2):pp. 198–215.

Smith, V. L. (1980). Experiments with a decentralized mechanism for public good decisions. *The American Economic Review*, 70(4):pp. 584–599.

Vanberg, C. (2008). Why do people keep their promises? an experimental test of two explanations1. *Econometrica*, 76(6):1467–1480.