Journal of Behavioral and Experimental Economics 53 (2014) 120–130

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Economics

journal homepage: www.elsevier.com/locate/jbee

# Can short-term incentives induce long-lasting cooperation? Results from a public-goods experiment☆

Lisa Bruttel[a,*], Tim Friehe[b,c]

[a] University of Potsdam, Department of Economics and Social Sciences, August-Bebel-Str. 89, 14482 Potsdam, Germany
[b] University of Bonn, Center for Advanced Studies in Law and Economics, Adenauerallee 24-42, 53113 Bonn, Germany
[c] CESifo, Munich, Germany

ARTICLE INFO

*Article history:*
Received 11 March 2014
Received in revised form 7 September 2014
Accepted 8 September 2014
Available online 17 September 2014

*JEL classification:*
C91
H26

*Keywords:*
Public-good game
Team
Punishment
Incentives
Experiment

ABSTRACT

This paper investigates whether providing strong cooperation incentives only at the outset of a group interaction spills over to later periods to ensure cooperation in the long run. We study a repeated linear public-good game with punishment opportunities and a parameter change after the first ten (of twenty) rounds. Our data shows that cooperation among subjects who had experienced a higher marginal return on public-good contributions or low punishment costs in rounds 1–10 rapidly deteriorated in rounds 11–20 once these incentives were removed, eventually trending below the level of cooperation in the control group. This suggests the possibility of temporary incentives backfiring in the long run. This paper ties in with the literature highlighting the potentially adverse effects of the use of incentives.



## 1. Introduction

### 1.1. Motivation and main results

This paper addresses the question of whether or not cooperation in the long run can be improved by the utilization of incentives that are strong only at the outset of a group interaction, that is, when a group of people has just been formed. In groups, many people are conditionally cooperative, that is, are ready to cooperate when they know or believe that others will do the same (e.g., Fischbacher and Gächter, 2010; Kocher et al., 2008). In many real-world circumstances, groups often fail to cooperate and find themselves in Pareto-dominated outcomes. In this study, we seek to shed light on whether it is possible to establish persistent cooperation by providing strong incentives only for the first few interactions made in the new group. The intuition for a finding in the affirmative would be as follows: strong initial incentives ease the attainment of the cooperative outcome in the beginning. Having played cooperatively several times, arriving at the cooperative outcome becomes natural even after the removal of strong incentives (given that the cooperative outcome clearly dominates the non-cooperative outcome in payoff terms even without strong incentives). Such a behavioral spillover from rounds with strong incentives to later interactions without strong incentives would rest on the influence of earlier rounds on subjects' beliefs about others' future behavior. In this way, our study is different from papers (to be discussed in detail in Section 1.2) that have explored the effectiveness of strong limited-term incentives provided only *after* play has converged to an inefficient outcome (Brandts and Cooper, 2006; Hamman, Rick and Weber, 2007) because, in our setting, temporary strong incentives can help build up trust – as high expectations are fulfilled by others – that can be capitalized on after the withdrawal of incentives.

For concreteness, imagine that a firm forms a team whose success is determined by the aggregate of unverifiable individual effort levels. In such a setup, no cooperation may result because team members bear the full costs of high effort but only reap full benefits

☆ We would like to thank Kate Bendrick, Gerald Eisenkopf, Urs Fischbacher, Katrin Schmelz, Simeon Schudy, Pascal Sulser, Verena Utikal, Irenaeus Wolff, Natalie Zimmer, an anonymous reviewer, and the editor Jeffrey Carpenter for very helpful suggestions.

* Corresponding author. Tel.: +49 3319774543.
*E-mail addresses:* lisa.bruttel@uni-potsdam.de (L. Bruttel), tim.friehe@uni-bonn.de (T. Friehe).

2214-8043/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.socec.2014.09.001

thereof if other team members also invest high effort. Accordingly, conditionally cooperative team members with the belief that other members will not cooperate will not supply effort. The firm may therefore attempt to induce high effort contributions (i.e., cooperation) from the start by implementing incentives facilitating cooperation over some initial time period. For example, the firm could put positive incentives into place by leveraging performance pay in relation to team success, or it might implement negative incentives by facilitating peer monitoring and reporting. If such incentives successfully induce cooperation between team members, their initial use may carry over to periods in which incentives are scaled back. In such a case, while incentives are in effect, they would induce cooperation between team members, and would additionally produce a cooperative spillover to team interactions in periods without strong incentives. However, the limited use of incentives facilitating cooperation may also have adverse effects, in the sense that team performance might become worse after their removal than it would have been without the initial use of these incentives. This could arise because team members perceive the reduction in rewards as unfair or because they lose intrinsic motivation, among other possible reasons (see, e.g., Falk and Kosfeld, 2006). Finally, the individual effort provision might be unrelated to the outcomes in earlier periods because past incentives are no longer relevant for current decisions.

In order to address the influence of short-run incentives on cooperation in the long run, we study a repeated linear public-good game with punishment opportunities in a fixed partners design. For such settings, the literature (see, for example, Fischbacher, Gächter and Fehr, 2001) has convincingly established that individuals' contributions are often positively correlated with what people expect others to contribute. As a result, earlier interactions in a setting with incentives facilitating cooperation may set the tone for interactions that occur after the removal of such incentives, due to their influence on participants' beliefs about the contributions of others.

Our central interest is with the long-run repercussions of using strong incentives only for the first few interactions among team members, an issue that has received relatively little attention in the literature. In principle, both low punishment costs and a high marginal return on public-good contributions are conducive to high individual public-good contributions, which is why we consider both as a possibility of setting incentives in our experiment.[1] Specifically, we consider rounds 11–20, contrasting the behavior of subjects who had a high marginal return on contributions to the public good in rounds 1–10 with the behavior of individuals who had a low marginal return on public-good contributions. Similarly, we contrast the behavior of participants who experienced low punishment costs in the first ten rounds with the behavior of subjects who had high punishment costs.

The experimental results suggest that the initial use of strong incentives is unable to ensure long-lasting cooperation. A high marginal return on public-good contributions very effectively increases contributions while it is in force. However, after this strong incentive is removed, cooperation rapidly deteriorates to levels that finally seem to be even smaller than those in the control group. Low punishment costs do not lead to significantly higher contributions while they are in effect (despite a significantly higher willingness to punish), but their phase-out results in similar reductions in cooperative behavior, possibly due to a significantly lower willingness to punish. Our findings are compatible with Hamman, Rick and Weber (2007) and conflicting with Brandts and Cooper (2006) in highlighting that temporary incentives have little prospect of inducing favorable long-lasting differences in behavior. Our results are in fact indicative of the possibility of backfiring, a feature observed in neither Hamman, Rick and Weber (2007) nor Brandts and Cooper (2006). In contrast to the decline of contributions in both treatments, our baseline scenario features average contributions that remain stable over time, thereby replicating the results of Nikiforakis and Normann (2008) who have used similar parameters.

### 1.2. Related literature

The present paper discusses the potentially adverse effects of using material incentives. In recent years, a number of published studies have warned about the dangers of unquestioned reliance on incentives (see Gneezy, Meier and Rey-Biel, 2011; Kamenica, 2012 for surveys). We are especially interested in the effects of strong incentives utilized during the initial phase of interactions, a relatively neglected aspect of this field.

In some studies, authors have been able to establish a link between previous conditions and later behavior. In a study without strategic interaction, Bruttel and Friehe (2014) establish that past levels of tax enforcement parameters (i.e., past levels of the sanction or the detection probability) are still relevant to the decision about tax evasion at later points in time. Similarly abstracting from strategic interactions, a field-experiment conducted by Meier (2007) establishes that matching donations in one time period can dampen charity contributions after the matching is phased out. Famously, Gneezy and Rustichini (2000) show that the introduction of fines for tardy parents at a child-care center increases tardiness even after the fines are abolished. In contrast, in the present paper, the interaction of past levels of important parameters (either the marginal return on public-good contributions or the punishment costs) and the strategic interaction between subjects takes center stage. In a strategic context, Gächter, Kessler and Königstein (2008) establish that the use of explicit incentives in an employment relationship can crowd out voluntary cooperation after the reliance on such incentives is discontinued.

We vary the experimental parameters in our treatment groups after ten rounds.[2] Similar intertemporal variation of central experimental parameters has been considered in Brandts and Cooper (2006) and Hamman, Rick and Weber (2007). We focus on whether it is possible to create long-lasting trust among group members by having an initial phase in which high expectations of participants are met by their group members. In contrast, Brandts and Cooper (2006) and Hamman, Rick and Weber (2007) are interested in overcoming a history of coordination failure by using temporary incentives once play has converged to an inefficient outcome. Specifically, Brandts and Cooper (2006) study fixed groups of four subjects in $3 \times 10$ rounds of a minimum-effort game designed to promote coordination failure over the first ten rounds. Inter alia, the authors contrast the behavior of subjects who experience a low

[1] Whereas the impact of a higher marginal return on contributions to the public good is direct, the effect of a variation of punishment costs is indirect. A lower level of punishment costs increases the willingness to punish low contributions, all else held equal; as a result, individuals are better off when they contribute more. There is empirical support for this hypothesized mechanism. For example, Carpenter (2007) has established that subjects in voluntary contribution mechanisms indeed demand more punishment when the price decreases, and Nikiforakis and Normann (2008) have shown that contributions increase monotonically with punishment effectiveness. See also the recent survey by Chaudhuri (2011).

[2] Instead of focusing on how behavior changes over time in a given game (i.e., exploring whether there are spillovers across periods of time) as we do in the present study, it is also possible to consider behavioral spillovers across games. For example, Cason, Savikhin and Sheremeta (2012) show that coordination in a minimum-effort game can be achieved through a behavioral spillover from a median-effort game when the former is played after the latter. Similar results have been found by Ahn et al. (2001) and Knez and Camerer (2000), among others.