Growth and Inequality in Public Good Provision *

Simon Gächter[†]
University of Nottingham,
CESifo and IZA

Friederike Mengel[‡]
University of Essex

Elias Tsakas[§]
Maastricht University

Alexander Vostroknutov[¶]
University of Trento

February 24, 2017

Abstract

In a novel experimental design we study public good games with dynamic interdependencies, where each agent's wealth at the end of period t serves as her endowment in $t+1$. In this setting growth and inequality arise endogenously allowing us to address new questions regarding their interplay and effect on cooperation. We find that amounts contributed are increasing over time even in the absence of punishment possibilities. Variation in wealth is substantial with the richest groups earning more than ten times what the poorest groups earn. Introducing the possibility of punishment does not increase wealth and in some cases even decreases it. In the presence of a punishment option inequality in early periods is strongly negatively correlated with group income in later periods, highlighting negative interaction effects between endogenous inequality and punishment.

*We thank Ingvild Almas, Benjamin Beranek, Tim Cason, Martin Dufwenberg, Tore Ellingsen, Drew Fudenberg, Santiago Oliveros, Ernesto Reuben and Marie-Claire Villeval as well as two anonymous Reviewers and participants in seminars and conferences at Essex, Maastricht, Lund (GIRLS 2014) and Toulouse (EEA ESEM 2014) for helpful comments. Friederike Mengel thanks the European Union (grant PERG08-GA-2010-277026) for financial support. Simon Gaechter gratefully acknowledges support from the European Research Council grant ERC-AdG 295707 COOPERATION and the ESRC network for Integrated Behavioural Science (NIBS, ES/K002201/1).

†School of Economics, University of Nottingham, University Park Campus, Nottingham NG7 2RD, United Kingdom.

‡Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom, e-mail: fr.mengel@gmail.com

§Department of Economics (AE1), Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands

¶Center for Mind/Brain Sciences, University of Trento, Via delle Regole 101, 38123 Mattarello, Italy.

1 Introduction

1.1 Motivation

Social dilemmas, where collective and private interests are in conflict, abound in economic and social life. Public good games have been used across disciplines as the standard tool to study a wide array of social dilemma situations. Those include joint ventures (Grossman and Shapiro, 1986), R&D cooperation (Cozzi, 1999; Kamien et al., 1992), political action funds of special interest groups or parties (Dawes et al., 1986), multilateral foreign aid or effort provision in work teams (Ostrom, 1990; Hamilton et al., 2003; Tirole, 1986). But also pricing or market sharing agreements by firms (Green and Porter, 1984) as well as many economic activities in the family (Becker, 1981) can be thought of as instances of cooperation that can be modeled with public good games. One feature that is common to many of these examples is that there are dynamic interdependencies: not only will the same set of people interact again, but previous outcomes affect future endowments (both in terms of the stock of physical and social capital).

In this paper we present a novel experimental design that captures such dynamic interdependencies. Our design builds on what has become the workhorse model to study public good provision in experiments (see e.g. Isaac et al. (1984); Andreoni (1995) or Fischbacher and Gächter (2010) among many others): participants are matched in fixed groups of four people to play the public good game for 10 or 15 periods. As in most other public good experiments we focus on the most challenging social dilemma situations, where the unique subgame perfect Nash equilibrium prescribes zero contributions by all group members, but where efficiency requires group members to contribute their entire endowment. We also conduct experiments where each group member can punish other group members by reducing their first stage earnings at a cost to themselves (Ostrom et al., 1992; Fehr and Gächter, 2000; Andreoni et al., 2003).

The key difference to previous research using this "standard" design is that each participant's wealth at the end of a period constitutes their endowment for the next period, whereas in the "standard design" endowments are allocated exogenously and tend to be the same in each round.[1] In our design, endowments are created endogenously, which leads to dynamic interdependencies.

We focus on two important implications of introducing these interdependencies. First, if overall contributions today are high, then there will be higher wealth in the next period (*growth*). Second, heterogeneity in contributions today creates *inequality* in endowments in the following period. Growth and inequality can interact with the possibility of punishment in different ways. The threat of punishment can lead to higher growth if it induces higher contributions, but punishment executed on the outcome path can induce a multiplier effect which can hamper growth severely. Maybe more interestingly, punishment can interact with inequality in non-trivial ways. In particular, rich group members can be largely "immune" to punishment by poorer group members, if the punishment that poor group members can afford is too small relative to the richest group member's wealth. As all endowment can be used for punishment, rich group members, on the other hand, might be able to punish others harshly at a relatively low cost to themselves. This asymmetry of punishment possibilities translates wealth inequality into inequality in power to punish. In addition, in unequal groups richer group members will typically be free-riders implying that punishment power could be in the "wrong hands". This raises the question of whether punishment will be as effective in increasing contributions and group income as it has been in settings without these dynamic interdependencies.

Our main findings can be summarized as follows. Even in the absence of punishment contri-

[1] Throughout the paper we will use the expression "standard design" to refer to the many studies of linear voluntary contribution experiments, where the one shot game is repeated for some number of periods. See Ledyard (1995) for a survey of the earlier literature and Chaudhury (2011) or Chapter 6 in Plott and Smith (2009) for a survey of more recent results.

bution and wealth levels display a strictly increasing trend over time. In terms of the realized potential for growth and the level of inequality, there is a lot of variation across groups. Individual earnings range between 2 Euros and 241 Euros. The Gini coefficient assumes the full range between 0 (equal wealth of all group members) and 1 (one group member appropriates the entire wealth) in the experiment.

Punishment (or the possibility thereof) does *not* increase wealth. This is true in both the 10 and 15-period variations despite the fact that people tend to contribute more in the treatment with punishment in the 10-period variations. We find evidence for two mechanisms behind this result: (i) in groups where inequality is high (above median) there is more anti-social than pro-social punishment, i.e. shirkers punish contributors more than vice versa and (ii) much of this punishment happens in early periods implying that resources are taken away exponentially.[2]

While the possibility of punishment does not increase wealth and in some cases even strictly decreases it, it also does not increase inequality on average. This is true despite the inequality-increasing presence of anti-social punishment. Analysis of data from Herrmann et al. (2008) shows that, in a comparable standard setting, punishment increases both wealth and inequality. In terms of the relationship between inequality and growth we find that inequality in period 2 is strongly negatively correlated with wealth in period 10 in the treatment with punishment possibilities. In particular, a 1% increase in inequality in period 2 leads to a $\approx 0.5\%$ decrease in wealth in period 10. Inequality and growth are positively related across groups with below median wealth and negatively related across groups with above median wealth.

In interpreting these results, it is important to note how the setting we introduce differs from the standard public good game described above. There are two main differences: (i) there is no consumption until the last round in our setting, i.e. one's entire wealth can be reinvested at the end of the period and (ii) endowments are endogenous, i.e. determined by previous outcomes. R+D cooperation often displays these features, but also the evolution of societies could be viewed under this lens. In the standard setting, by contrast, there is full consumption, i.e. no wealth can be reinvested and endowments are exogenous (and stationary). Volunteering, e.g. at a food bank, seems a good example falling into this category. Other types of volunteering, such as in natural conservation or archiving, are examples involving stationary exogenous endowments, but no or little consumption. The case of full consumption with endogenous endowments describes the one-shot game. Finally, note that many applications, such as infrastructure investments, multilateral foreign aid or pricing agreements will fall somewhere in between these extremes with some, but not full consumption and with partially endogenous endowments.

1.2 Literature Review

To our knowledge our experiment is the first to study public good provision with dynamic interdependencies and endogenously arising asymmetric punishment possibilities.[3] As such it contributes to studies of public goods with dynamic interdependencies more broadly. Those include Battaglini et al. (2014) who study the Markov perfect equilibrium dynamics in the provision of a durable public good over time where there is consumption in each period. They find evidence of significant under-provision relative to the interior equilibrium. Duffy et al. (2007) studied

[2] History books report many examples where "shirking group members" have assumed power by exploiting asymmetric punishment opportunities. Standard Oil reportedly sent out thugs to raid the premises of competitors as a form of punishment (Josephson, 1962). Adler (1985) discusses endogenously arising asymmetries in punishment possibilities in a study of upper-level drug-dealing and Johnson and Earle (1987) among North-American Indians.

[3] In an unpublished Master thesis Huck (2006) has conducted a dynamic public good game without the possibility of punishment and related contributions to personality characteristics elicited in a questionnaire. He also did not analyze growth and inequality but, like us, he finds a pattern of increasing contributions (in terms of absolute amounts) and no endgame effect.

threshold public good games with multiple contribution rounds, where, theoretically, "completion equilibria" (with positive contributions) do exist. They find that, as in the standard setting, contributions do decline over time (see also Croson and Marks (1998) among others). Noussair and Soo (2008) and Cadigan et al. (2011) study dynamic public good settings where the current return from the public good depends on past contributions. Other studies link public good games over time via explicit reputation mechanisms (e.g. Milinski et al., 2002). Guererk et al. (2017) have recently studied a setting where games are linked via endowments. They use a non-linear exchange rate, though, which effectively eliminates the possibility of exponential growth and contains inequality. Consequently their results are more similar to those obtained in the standard setting.

Our results also contribute to research on the impact of inequality on public good provision. Most existing literature has studied the effects of *exogenous* income inequality. Chan et al. (1996) experimentally test a prediction by Bergstrom et al. (1986), where public good provision increases with inequality in the income distribution in an equilibrium with positive contributions. They find that group behavior conforms with the theoretical prediction. Other authors have found that exogenous income inequality decreases contributions (van Dijk et al., 2002; Ostrom et al., 1994) or found no effect (Chan et al., 1999). Reuben and Riedl (2013) find that without punishment there is no effect of income inequality on contributions, while with punishment participants contribute proportionally to their endowments, leading to increased contributions. All these papers deal with one-shot games and *exogenously* imposed inequality. Sadrieh and Verbon (2006) study exogenous inequality in the Bergstrom et al. (1986) setting with the possibility of growth. They find that exogenous variations of inequality are mostly neutral to growth.[4] The latter result contrasts with our finding that endogenous inequality is negatively related to growth in the presence of punishment. One possible reason for the difference between these results is that, as discussed above, with endogenous inequality the power to punish tends to lie "in the wrong hands" (those of free-riders).

Our paper also contributes to the literature on the standard setting of repeated public good games surveyed in Chaudhury (2011). One insight emerging from this literature is that societies can maintain high levels of contributions and wealth if there is a threat of punishment to free-riders (Ostrom et al., 1992; Fehr and Gächter, 2000, 2002; Andreoni et al., 2003).[5] While some studies have found that punishment can lead to lower payoffs if the horizon of the game is short (Egas and Riedl, 2008; Dreber et al., 2008), under a long horizon punishment in the standard setting has been found to be strictly beneficial (Gächter et al., 2008). Punishment does not increase wealth in our setting neither under a 10-period nor a 15-period horizon. We do not study longer horizons, but since many groups in the treatment with punishment destroy (almost) all wealth by period 10, these groups will be unable to recover even with a very long horizon. Some studies of the standard setting investigate how an increased marginal per capita rate of return affects contributions (Isaac and Walker, 1988; Goeree et al., 2002). There is some relation of these studies to our setting, since the possibilities of exponential growth provides increased incentives to contribute.

Finally, the dynamic setting we study also relates to other dynamic games studied in economics. In the common pool resource game (CPR, Ostrom et al. (1994)) players extract from a resource in each period with the non-extracted part growing at a fixed rate. There are various differences between this setting and our dynamic public good game (apart from the frame). One is that whatever is contributed in our public good game is mechanically shared among all participants. How the benefits of non-extraction are shared depends on players' strategies. On the

[4] In a trust game with limited growth Greiner et al. (2011), however, find that both exogenous and endogenous variation in income affect growth.

[5] Cason and Gangadharan (2015) note, though, that the impact of punishment is weaker in nonlinear social dilemmas.

other hand what is withheld in our public good game in one period can be contributed in the next. What is extracted in the CPR game, however, cannot be reverted to the pool in future periods. As a consequence of these various differences, both games have different equilibria. The CPR game has an interior stationary Markov equilibrium where players extract too much relative to the efficient benchmark (Mailath and Samuelson, 2005). By contrast, in our dynamic public good game the only equilibrium has zero contributions in all rounds (Section 3).

Herr et al. (1997) study two variations of the CPR game, one where extraction by one player increases the cost of extraction by others in the current period only, and one where extraction increases the costs for all future periods. While both versions have interior equilibria, the latter type of externality should worsen the CPR problem, i.e. lead to lower efficiency. Experimental behaviour is in line with equilibrium benchmarks. Botelho et al. (2014) study a game that lies in between a classic CPR and Centipede game, in that they study a CPR game which ends as soon as the resource stock falls below a certain level. They show that increasing uncertainty about that level increases extraction and hence the inefficiency. Battaglini et al. (2014) discussed above can be viewed as a public good study within the CPR paradigm.

This paper is organized as follows. In section 2 we present the experimental design. Section 3 discusses the theoretical predictions and summarizes our research questions. Section 4 contains our main results. In Section 5 we discuss the mechanisms underlying these results. Section 6 concludes. An Online Appendix contains screenshots, experimental instructions and the questionnaire, proofs of the theoretical predictions and several additional results, tables and figures.

2 Experimental Design

In our experiment participants play a sequence of dynamically interdependent public good games in two main treatments: (1) treatment NOPUNISH in which participants only play the public good game; (2) treatment PUNISH where, after each period, participants can also subtract tokens from other members of their group at a cost. We describe these treatments in turn.

Treatment NOPUNISH At the beginning of the experiment participants are randomly matched into groups of 4, which stay the same throughout the experiment. Participants are indexed $i \in I = \{1, 2, 3, 4\}$. Before period 1, each participant is given 20 tokens as initial endowment. In each period participants can divide their tokens into two accounts: a *private* account and a *group* account. The private account, which has a return of 1, is for their personal use only and other participants cannot influence or benefit from the tokens in the private account. The group account is different: all 4 participants in the group can place their tokens into the group account. The group account has a return of 1.5 and after the return is calculated, the tokens put into the group account are equally divided among all 4 participants. Therefore, if participant i has N_i^t tokens *before* period t and she places c_i^t tokens into the group account, then the rest of the tokens automatically go to the private account and at the *end* of period t, she will have $N_i^t - c_i^t + \frac{1.5}{4}\sum_{j=1..4} c_j^t$ tokens. In the NOPUNISH treatment, this amount also corresponds to the number of tokens *before* period $t+1$ (the endowment at $t+1$):

$$N_i^{t+1} = N_i^t - c_i^t + \frac{1.5}{4}\sum_{j=1}^{4} c_j^t.$$

Therefore, in period $t+1$, the number of tokens that each participant can invest depends on the choices of all group members in previous periods (endogenous endowments). This is what makes our set-up different from the standard setting where the amount of tokens before each period is fixed and the earnings from previous periods cannot be used for investment into the public good (exogenous stationary endowments). At the end of each period participants observe information about endowments and contributions of all group members (Figure A.1 in Online Appendix A).

Treatment PUNISH The PUNISH treatment is the same as the NOPUNISH treatment with the difference that, after all participants in the group have allocated their tokens and observed others' allocations (as in NOPUNISH), they have a possibility to subtract tokens from individual members of the group. The cost of subtraction is $\frac{1}{3}$ of the tokens subtracted.

More specifically, let $p_{i,j}^t$ denote the cost of punishment that participant i incurs after punishing participant j by $3p_{i,j}^t$ tokens at period t, and denote by $W_i^t := N_i^t - c_i^t + \frac{1.5}{4}\sum_{j=1}^4 c_j^t$ the amount of pre-punishment tokens of player i. The amount of tokens that participant i has after punishment and hence at the start of period $t+1$ is then

$$N_i^{t+1} = \min_{J \in \mathcal{J}_i^t} \left\{ W_i^t - \sum_{j \in J} 3p_{j,i}^t \right\} - \sum_{j \neq i} p_{i,j}^t$$

where $\mathcal{J}_i^t$ contains all nonempty subsets $J \subseteq I$ such that (a) $i \notin J$, (b) if $k \in J$ then $j \in J$ for all $j \in \{1, \ldots, k\} \setminus \{i\}$, and (c) $W_i^t - \sum_{j \in J} 3p_{j,i}^t \geq 0$. This formulation accounts for the fact that punishments which would set a player's wealth below zero are not executed.[6] After the punishment phase each participant observes the following information before the next period starts (see Figure A.3 in Online Appendix A): the amount of tokens that each other member of the group subtracted from her; the total amount of tokens subtracted from each member of the group; the amount of tokens each group member has at the end of the period, which serves as the endowment in the next period. Labels for group members on these screens are randomized across periods.

Additional Treatments As we outlined above the dynamic interdependencies in this setting create two types of effects: (i) they create the possibility for endogenous growth and (ii) they create endogenous inequality and hence asymmetries in the power to punish others. To better understand the impact of these two forces we ran additional treatments, where we artificially eliminate growth (treatments NOPUNISH-NOGROWTH and PUNISH-NOGROWTH) but keep endogenous inequality or where we artificially eliminate inequality but keep growth (treatments NOPUNISH-NOINEQUALITY and PUNISH-NOINEQUALITY). We will discuss these treatments in more detail in Section 5.

Length variations In addition to the punishment variation, we added a second treatment variation which changes the number of repetitions of the public good game. We conducted sessions with 10 periods and sessions with 15 periods to understand how a longer horizon might affect behaviour and hence get some insights into how participants might perceive this environment strategically.[7] Due to the (potentially) exponential increase of earnings over time in our setting,

[6] If participants' punishment plans are infeasible, e.g. because in the aggregate they would set a player's wealth below zero, then only player 1 and 2's plans are executed if feasible (as labeled in the software). If those are also infeasible then only player 1's plan is executed. Since subject labels are randomized by the software this does not create imbalances between subjects. Note also that all players have to pay p_i^j for all punishments intended irrespective of whether they are feasible and of whether punishing sets their own payoffs to zero. Since participants' earnings can still become negative if they punish someone else, this restriction does not affect the set of subgame perfect Nash equilibria. In imposing this restriction we follow the literature on the standard setting (see e.g. Fehr and Gächter (2000)).

[7] Gächter et al. (2008) have found that the length of the horizon can affect contributions in the standard setting, particularly in treatments with punishment.

we limited the long horizon to 15 periods.[8] Participants were informed about all these details of the design at the beginning of the experiment.

Other Details Participants answered a sequence of questions at the end of the experiment (see Online Appendix B). We report summary statistics on these characteristics in Tables 14 and 15 in Online Appendix F, where analysis of the questionnaire data can be found. In all treatments the amount that participants received at the end of the experiment was equal to a €2 show up fee plus the amount of tokens after the last period converted into Euros. 1 token was equal to 0.05 Euros. All experiments were run at the BEE-Lab at Maastricht University and used z-tree (Fischbacher, 2007). Table 8 in Appendix D shows the exact order of sessions for the main treatments. In total 656 participants took part in our experiment: 152 in treatment NOPUNISH (38 groups), 144 in treatment PUNISH (36 groups) and the remainder in one of the additional treatments. Table 7 in Appendix D summarizes the treatment structure and the number of independent observations, participants and sessions for each treatment. Average earnings were 11.94 Euros with a minimum of 2 Euros and a maximum of 241.65 Euros. No other sessions apart from those reported here were conducted and there were no pilot studies.

3 Theoretical Background and Research Questions

In this section we briefly discuss the theoretical predictions and then present our research questions. We focus on the standard solution concepts for dynamic games with observable actions, i.e., Nash equilibrium (NE) and subgame perfect equilibrium (SPE) under the standard textbook assumptions of self-interest and common knowledge of rationality. First, observe that the games that describe our treatments are not repeated games, as the set of available actions at each subgame depends on the moves that have been already realized. Therefore, unlike in the standard setting, we cannot directly apply the wide range of well-known results on finitely repeated games. Nevertheless, it turns out that SPE leads to very similar predictions as in the standard setting. On the equilibrium path all players contribute 0 in all periods, and moreover in the punishment treatment no player ever punishes. The purpose of stating these results is not to argue that they will be good predictions of behaviour, but rather to give the reader a sense of the incentive properties of the public good game with dynamic interdependencies. In this section we will state the result informally. Formal statements and proofs can be found in Online Appendix C.

Proposition 1 *Consider the public good game with growth as defined above. The unique SPE of both the game without punishment and with punishment is such that all players contribute 0 throughout the entire game (both on and off the equilibrium path).*

As is the case with the standard setting, positive contributions can be sustained under different assumptions. Particularly reputational models that invoke the existence of behavioural types such as e.g. Kreps et al. (1982) have the potential to induce positive contributions in either setting, but also to drive a wedge between the two settings in terms of the extent of cooperation found. More precisely, following Kreps et al. (1982), let us assume that there are two types of agents: (i) "standard" rational and self interested types and (ii) conditional cooperators who start out by contributing 10 tokens and then match at each time t the minimal contribution in their group at $t-1$. We assume that all agents are of type (i) and that each agent has a minimal doubt that other agents may be of type (ii), i.e. conditional cooperators. More precisely we denote by μ the

[8] There was one session that was intended to be 15 periods, but where a computer crashed in period 11. We added this session as a 10 period session to the data, but none of the primary findings are affected if we exclude this session.

probability with which agents believe that all other agents are conditional cooperators. We also denote by T the length of the game, which in our experiment is either 10 or 15 periods. We can then state the following result.

Proposition 2 *Assume $\mu < \frac{10}{35}$. All sequential equilibria involve positive contributions on the outcome path*

- *if and only if $\mu > \hat{\mu}_S := \frac{50}{20T+5}$ in the standard setting and*
- *if and only if $\mu > \hat{\mu}_G := \frac{6.25}{4 \cdot 1.5^T - 2.25}$ in our setting.*

For any fixed $\mu < \frac{10}{35}$, the equilibrium always involves strictly more periods of positive contributions in our setting.

The proof of this result can be found in Online Appendix C. Because of the incentives provided by exponential growth, even a small doubt that some agents may be conditional cooperators suffices to induce positive contributions on the equilibrium path. In the standard setting, by contrast, such doubts have to be substantial ($\mu > 0.19$) to even get one period of positive contributions. In fact, if $0.03 < \mu < 0.19$ and $T = 10$, then positive contributions can only be sustained in the setting with growth. For any given μ, more periods of contributions will be observed in the setting with growth. And, across the periods where contributions are positive, the amount contributed is constant in the standard setting and increasing in the setting with growth (for a related result in Centipede games, see McKelvey and Palfrey (1992)).

We next summarize our research questions. First and foremost, we are interested in whether participants contribute positive amounts and in how much wealth is created in this game.

Q1 *Are there sustained positive contributions in our setting? How much wealth is created?*

Apart from wealth creation, we are also interested in the amount of inequality created both within and across groups. Of particular interest is the relation between inequality and growth. The sign of the relation between inequality and growth of societies as well as the causal link between the two has been at the center of a debate in macroeconomics and development (Barro, 2000; Forbes, 2000; Persson and Tabellini, 1991). Hence, while Q1 asks whether societies can see positive contributions and increasing wealth over time, our next question Q2 asks how much inequality is generated in this process and how inequality affects growth. Our setting allows us to address this question both within and across groups. In particular, we ask whether "rich" or "poor" groups will be more unequal and how initial inequality affects growth and hence final wealth.

Q2 *How much inequality is generated in our setting? What are the consequences of endogenous inequality for growth?*

Last, we are interested in the effects of punishment in our setting. Punishment has been shown to be effective in the standard setting in securing high contributions and wealth (Gächter et al., 2008). In our setting, however, punishment does destroy resources which could otherwise be used productively in the following period. Punishment in the dynamic setting can have additional adverse effects that operate via endogenously created inequality. Because earnings are carried over across periods, free-riders in our setting will have more resources to punish others than contributors. Hence, the possibility of punishment could strengthen existing inequalities because it makes shirking individuals more powerful. This in turn could undermine the effectiveness of punishment.

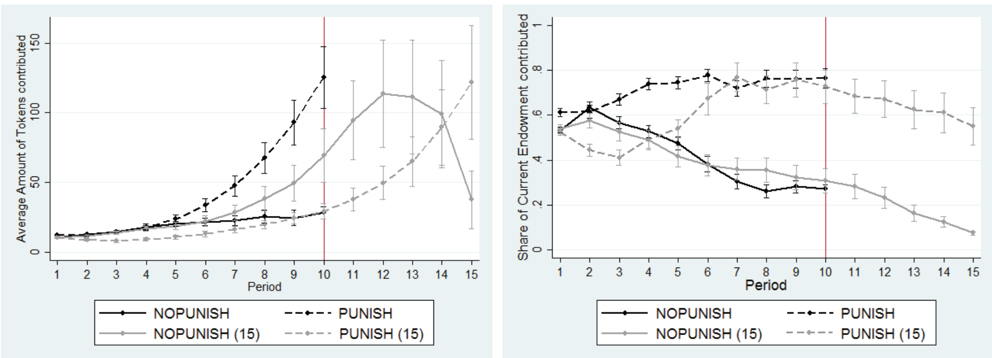
Q3 *How does the possibility of punishment affect contributions, growth and inequality?*

4 Results

This section comprises our main results. Subsection 4.1 focuses on Question 1, i.e. on how much participants contribute and how much wealth is generated in this setting. Subsection 4.2 focuses on Question 2, i.e. how much inequality is endogenously created. We study the effect of punishment (Q3) within each of these subsections. In Section 5 we discuss additional results and possible mechanisms behind our results.

4.1 Provision of the Public Good and Wealth Creation

We start by discussing contributions. Panel (a) in Figure 1 shows the average amount of tokens participants contributed over time. Contributions are clearly non-zero and are increasing over time in all treatments. Even without punishment participants contribute about 10 tokens in the first period and then steadily increase this amount over time. Contributions flatten out towards the end of the experiment and there is even a statistically significant drop in period 15 in the long horizon game. Note that increasing contributions over time imply that participants have increasing endowments over time. Hence increasing contributions do not necessarily imply that participants contribute increasing shares of their endowments. Panel (b) in Figure 1 shows the share of overall endowments contributed over time in all main treatments. In NOPUNISH participants contribute around 55% of their endowment in period 1. This amount steadily decreases between periods 2-8 and is roughly constant afterwards at a level of about 35% of endowments. In the 15-period games shares follow a similar pattern until period 10, then start to decrease again afterwards.



(a) Amounts (b) Share of Endowment

Figure 1: The average amount of tokens contributed over time in treatments NOPUNISH (Without Punishment) and PUNISH (With Punishment) in the 10-period and 15-period variations. Bars indicate one standard error of the mean.

Panel (a) in Figure 1 presents a stark contrast to typical 10-period public good games in the standard setting, where contributions are decreasing over time (see e.g. Figure 3 in (Fehr and Gächter, 2000)).[9] It should be kept in mind, though, that in the standard setting with stationary exogenous endowments, the dynamics of contributions over time is the same irrespective of whether it is measured in absolute terms or as a share of endowment. In Fehr and Gächter (2000), for example, contributions equal around 50-60 percent of endowments in Period 1 and decrease to around 10 percent of endowments by Period 10. Panel (b) in Figure 1 shows that in our setting contributions stabilize at a higher level ($\approx 35\%$) in the 10-period games even if

[9]Decreasing patterns in public good games are found for MPCR's roughly in the range of the MPCR of our one shot game of 0.375. For higher MPCR's (e.g. 0.75) contributions are still decreasing, but more weakly so (Isaac and Walker, 1988; Zelmer, 2003).

measured as share of endowments contributed. One possible explanation is that the dynamic incentives with the possibility of exponential growth have a similar effect to that of increasing MPCR's in the standard setting (Isaac and Walker, 1988; Goeree et al., 2002).[10] This higher level of shares contributed seems not to be sustainable in the longer run, though. By period 15 a similar share of endowments is contributed as in the standard setting after 10 periods (10-20%).

	Wealth		*Ranksum Test*	
	NOPUNISH	PUNISH	Higher Ranksum	p-value
$t = 10$ Mean	439.69	540.30	NOPUNISH	0.0928
$t = 10$ Median	299	143		
$t = 10$ Std.Dev.	345.39	793.17		
$t = 10$ Max	1663	2724		
$t = 10$ Min	158	0		
Observations	23	21		
$t = 15$ Mean	1503.46	713.06	NOPUNISH	0.3720
$t = 15$ Median	422	248		
$t = 15$ Std.Dev.	2873.79	1279.54		
$t = 15$ Max	8687	5096		
$t = 15$ Min	155	0		
Observations	15	15		

Table 1: Summary Statistics on variable **wealth**.

Possibly of more interest are the implications contributions have for wealth generation and growth. To measure growth we define a variable "wealth" which sums the endowments of all participants in a given group at the beginning of the following period, i.e. wealth in t would be $\sum_{i \in I} N_i^{t+1}$ using the notation from Section 3. Before the start of period 1 wealth will be 80 in all groups by construction. The maximal wealth that can be reached in period 10 (if everyone contributes their entire endowment in each period) is approximately 3075 tokens or €153. After period 15 the maximum is approximately 35000 tokens or €1751. The minimal wealth that can be reached (if no one ever contributes anything) is 80 in NOPUNISH and negative in PUNISH (if someone punishes others and is heavily punished in the same period). Table 1 shows some summary statistics regarding wealth. Groups do achieve growth on average. While there is clearly growth, groups do not realize the maximal potential efficiency. In the 10-period games groups reach on average a level of 439.69 tokens out of 3075 maximally possible or 14.29%. There is large heterogeneity with the richest group reaching 1663 tokens and hence more than 10 times more than the poorest group.

Figure 2 shows the dynamics of wealth over time. Panel (a) focuses on all groups, panel (b) on those with above median wealth in Period 10 ("successful" groups) and panel (c) on those with below median wealth in Period 10 ("unsuccessful" groups). Average wealth is increasing across periods (see coefficients β_1 and β_4 in Tables 9 and 10 in Appendix D) and is substantially above 80 once period 10 is reached (Table 1). Not all groups achieve growth, however. There are several groups where wealth does not rise (substantially) above 80. Figures E.1-E.2 in Online Appendix E illustrate wealth (and Gini coefficients) over time in different matching groups.

Result 1: Contributions and Wealth creation *Amounts contributed are positive and increasing over time even without punishment. Wealth is growing over time, but there is large variation across groups with the richest group earning more than 10 times more than the poorest after 10 periods and more than 55 times more after 15 periods.*

[10]In a CPR game Herr et al. (1997) observe a pattern of relatively constant extraction over time. They find extraction rates that are close to the prediction of the interior subgame perfect Nash equilibrium. In our setting, the SPE is not interior, but prescribes zero contributions instead, while the contribution levels we find experimentally are substantially above that. Kimbrough and Vostroknutov (2015) study a dynamic CPR game where for sufficiently high regrowth rates the resource will be provided indefinitely in equilibrium.

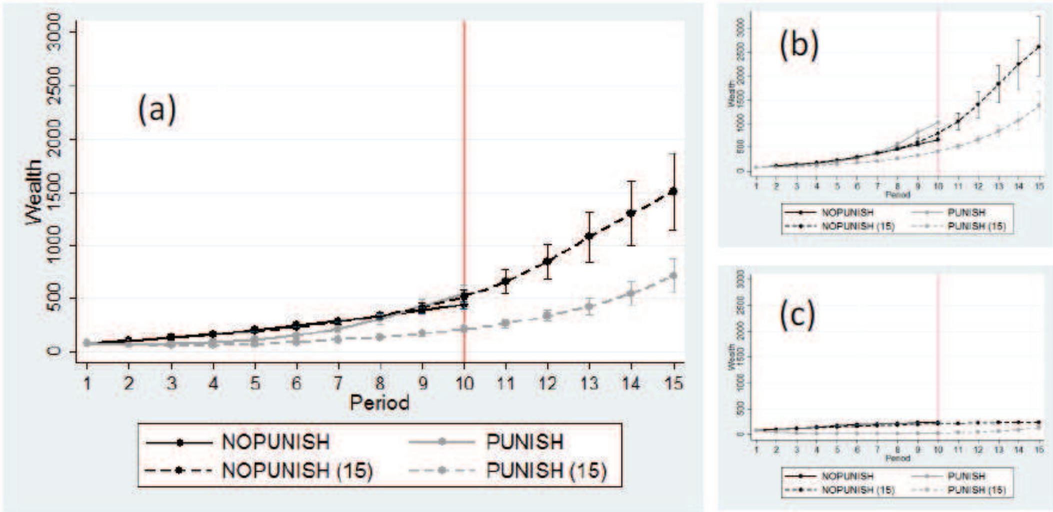


Figure 2: Average wealth over time across treatments. Panel (a) all groups, (b) groups with above median wealth and (c) groups with below median wealth. Bars indicate one standard error of the mean.

	Wealth					
		10 period games			*15 period games*	
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	Below Median	Above Median	All	Below Median	Above Median
PUNISH	107.0	-213.0***	361.3	-790.4**	-101.3	-1,239
	(89.2)	(17.68)	(280.2)	(395.6)	(101.3)	(1,408)
Constant	439.7***	238.3***	659.4***	1,503**	239.7***	2,609*
	(71.46)	(12.51)	(117.9)	(732.2)	(23.86)	(1,269)
Observations	44	22	22	30	15	15
R-squared	0.008	0.873	0.074	0.033	0.061	0.049

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2: OLS regression of wealth on treatment dummy. Period 10 (15) data only.

Effects of Punishment on Contributions and Wealth

We next consider the effect of punishment on contributions and wealth. Absolute contributions end up higher in PUNISH compared to NOPUNISH (Panel (a) in Figure 1), but they start to differ only from period 7 onwards in the 10-period games and in the 15-period games they are higher only in the very last period. In terms of shares contributed (Panel (b) of Figure 1), in PUNISH contributions are roughly constant across all periods at a level of about 60%, though there is a slight decrease towards the end of the 15 period games. This is above the share at which groups stabilize in treatments without punishment.

However, groups in PUNISH are also poorer. Median wealth is higher in NOPUNISH compared to PUNISH (Table 1) and the difference in mean ranks is statistically significant in the 10-period games according to a one-sided ranksum test ($p < 0.0001$ for all periods; $p = 0.0928$ for $t = 10$ data only). To assess the statistical significance of differences in means we run OLS regressions where we regress wealth on a treatment dummy for PUNISH (Table 2). These regressions show that differences in means are only significant for below median groups in the 10-period games which earn on average 213 tokens less in PUNISH compared to NOPUNISH. Several of these groups, in fact, end up with zero income in PUNISH. In the 15-period games differences are statistically significant across all groups with mean wealth in period 15 being about 790 tokens lower in PUNISH compared to NOPUNISH. The fact that punishment seems more harmful the longer the horizon stands in contrast to results obtained in the standard setting by e.g. Gächter et al. (2008). Columns (1) and (2) of Table 9 in Appendix D assess differences in time trends using a linear (column (1)) and square (column (2)) polynomial in period. While wealth is increasing in both treatments ($\beta_1, \beta_1 + \beta_3$), the linear trend in column (1) is no different across treatments (β_3). The square polynomial does reveal differences in time trends, though. While wealth is initially increasing at a lower rate in PUNISH (β_3), it increases at a faster rate in later period (β_5). Figure 2 illustrates these results. It shows that wealth is lower in PUNISH compared to NOPUNISH across all periods except for the last three periods in the 10-period games.

Overall, these results indicate that the threat of punishment is not needed to sustain high contributions in public good games with growth and does not increase wealth. Instead the dynamic evolution of the public good drives contributions and wealth. This is in stark contrast to what we know from other social dilemma games, where punishment has been shown to be successful in both raising contributions (in Fehr and Gächter (2000)'s study on the standard setting by around 600%) as well as wealth *in later periods* (in Fehr and Gächter (2000)'s study from period 4 onwards (Result 8 in Fehr and Gächter (2000)). In Section 5 we discuss in more detail why punishment is less effective in our setting. There we also analyze data from an experiment by Herrmann et al. (2008) conducted in the standard setting and show that in this case punishment leads to unambiguously higher wealth.

Result 2: Effect of Punishment on Wealth creation *The possibility of punishment does not increase wealth. In the long horizon games (15 periods), average wealth is even lower in PUNISH compared to NOPUNISH.*

4.2 Inequality

In this subsection we focus on the amount of inequality created endogenously in our setting. As measure of inequality we use the Gini coefficient as defined in Deaton (1997). The smallest possible value the Gini coefficient takes is zero (if all four group members own one fourth of the wealth) and the largest possible value it takes is one (if one group member holds the entire wealth). Table 1 shows some summary statistics regarding the Gini coefficient. The period 10 Gini coefficient ranges between 0 and 0.43 in NOPUNISH with a median of 0.22 and assumes the full range between 0 and 1 in PUNISH with a median of 0.03. The period 15 Gini coefficient ranges between 0.03 and 0.49 in NOPUNISH and between 0 and 0.67 in PUNISH.

	Gini coefficient		*Ranksum test*	
	NOPUNISH	PUNISH	Higher Ranksum	p-value
$t = 10$ Mean	0.21	0.22		
$t = 10$ Median	0.22	0.03		
$t = 10$ Std.Dev.	0.12	0.35	NOPUNISH	0.0395
$t = 10$ Max	0.43	1		
$t = 10$ Min	0.00	0		
Observations	23	21		
$t = 15$ Mean	0.18	0.14		
$t = 15$ Median	0.10	0.05		
$t = 15$ Std.Dev.	0.16	0.20	NOPUNISH	0.1508
$t = 15$ Max	0.49	0.67		
$t = 15$ Min	0.03	0		
Observations	15	15		

Table 3: Summary Statistics on the variable **Gini coefficient**. Period 10 (15) data only.

Figure 3 illustrates the dynamics of the Gini coefficient over time. The figure shows that in NOPUNISH inequality is sharply increasing from the initial value of zero and then increases more slowly between periods 2-10 to reach a level of ≈ 0.2 in period 10 (β_2 and β_4 in column (2) of Tables 11 and 12 in Appendix D). Interestingly, these patterns almost perfectly mimic the trends in inequality identified in a recent paper by Nishi et al. (2015) (see their Figure 2) who study

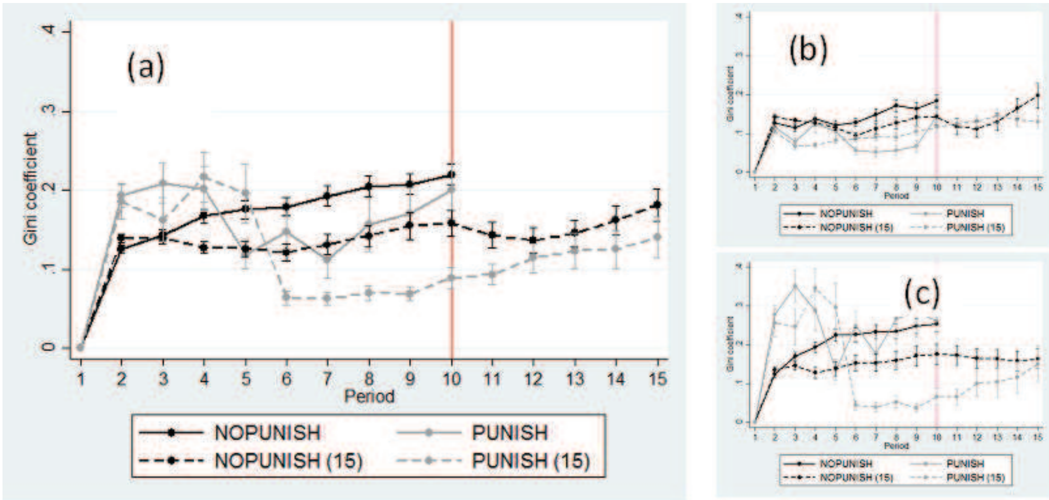


Figure 3: Average Gini coefficient over time across treatments. Panel (a) all groups, (b) groups with above median wealth and (c) groups with below median wealth. Bars indicate one standard error of the mean.

the effect of visibility of wealth (and inequality) in a cooperation game played on a network. A longer horizon seems to lead to slower growth in inequality, but by period 15 Gini coefficients are not statistically different from those reached in period 10 of the shorter horizon games. Panel (b) shows groups with above median wealth, which display lower levels of inequality than those with below median wealth illustrated in Panel (c) (ranksum test $p < 0.001$ all periods, $p = 0.1096$ period 10, $p > 0.1$ period 15).

	Gini coefficient					
		10 period games			*15 period games*	
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	Below Median	Above Median	All	Below Median	Above Median
PUNISH	-0.020	0.008	-0.040	-0.041	-0.013	-0.068
	(0.035)	(0.128)	(0.062)	(0.033)	(0.110)	(0.072)
Constant	0.220***	0.253***	0.184***	0.181***	0.163***	0.197***
	(0.024)	(0.038)	(0.029)	(0.023)	(0.053)	(0.064)
Observations	44	22	22	30	15	15
R-squared	0.002	0.000	0.019	0.013	0.001	0.055

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: OLS regression of Gini coefficient on treatment dummy. Period 10 data only.

Result 3: Inequality. *Mean Gini coefficients are increasing over time reaching* ≈ 0.2 *in the last period. There is substantial variation in Gini coefficients across groups, particularly in PUNISH where the Gini assumes the full range between 0 and 1 in the 10-period games.*

Effects of Punishment on Inequality

We next study the effect of punishment on inequality. Mean Gini coefficients are similar across treatments and there are no statistically significant differences in mean Gini coefficients between NOPUNISH and PUNISH in period 10 or 15, respectively (see Table 4). By contrast, analysis conducted with data from a standard public good game in Section 5.4 shows that in this setting punishment leads to more inequality. In fact, in our case the median Gini coefficient tends to be even higher in NOPUNISH compared to PUNISH (Table 3). The difference in mean ranks is statistically significant (one-sided ranksum test $p = 0.0395$) only in the 10-period games. There is also more variation in PUNISH (one-sided variance ratio test on period 10 data only, $p < 0.0001$; period 15 data: $p < 0.0001$) where the Gini coefficient assumes the full range between 0 and 1 in the 10-period games and between 0 and 0.67 in the 15-period games.

Figure 3 illustrates these results as well as differences in time trends. Across all groups (Panel (a)) the dynamics of inequality seem similar across treatments with somewhat more volatility in PUNISH. Columns (1)-(2) in Table 11 show that there are few statistically significant differences in these time trends. The dynamics of inequality over time seem also quite similar in above median groups, where the difference is mostly one of means (Panel (b)). In below median groups, however, time trends are quite different across treatments (Panel (c)). In NOPUNISH inequality is steadily increasing over time at a slow rate, while in PUNISH the Gini coefficient seems to follow a different pattern. After increasing initially, it decreases sharply around periods 4-7 and then starts to increase again. One possible interpretation is that this pattern reflects cycles of reciprocal punishment. Depending on who punishes (shirkers or high contributors) inequality increases or decreases. We discuss differences in punishment behaviour between below and above median groups in more detail in Section 5.3.

Result 4: Effect of Punishment on Inequality. *Mean inequality (Gini) is not statistically different in PUNISH compared to NOPUNISH in period 10 (15). There is more across-group variation in inequality in PUNISH.*

5 Discussion and Additional Results

In this section we discuss some of the potential mechanisms underlying our main results. Section 5.1 discusses results on the relationship between growth and inequality. In Section 5.2. we will try to tease apart the effect of exponential growth opportunities from the effect of endogenously created inequality. In Section 5.3, we provide some additional results on punishment, which should help us get a deeper understanding of why punishment is less effective here than in the standard setting. Finally, in Section 5.4 we use data from experiments conducted by Herrmann et al. (2008) to study wealth and inequality in the standard setting.

5.1 The Relation between Growth and Inequality

We study the relationship between growth and inequality across and within groups starting with across-group inequality. In NOPUNISH the two variables are essentially uncorrelated. The Spearman correlation coefficient is -0.1840 in above median groups and -0.0630 in below median groups, neither of which is statistically different from zero. In PUNISH, by contrast, there is a substantial and statistically significant relationship between these two variables. In above median groups the correlation is negative (Spearman correlation coefficient $\rho = -0.6091^{**}$): higher wealth comes with lower inequality. In fact, most groups that achieve very high levels of wealth in period 10, have a Gini coefficient of almost zero. For below median groups, the correlation is positive ($\rho = 0.4683^{*}$): higher wealth comes with higher inequality. As a consequence inequality is highest in groups with intermediate levels of wealth in PUNISH. Figure D.1 in Appendix D illustrates the correlation between period 10 wealth and Gini coefficients for both treatments. The same patterns hold in the 15-period games, but, possibly due to lower power, there is no statistical significance.

These differences between above and below median groups appear already as early as in the second period of the game. For NOPUNISH both successful (above median wealth in period 10) and unsuccessful (below median wealth in period 10) groups have a Gini coefficient of 0.13 on average in period 2. In the 15-period games these numbers are 0.13 and 0.14, respectively. In PUNISH, on the other hand, there are substantial differences. Groups that are eventually successful have a Gini coefficient of about 0.11 in period 2 (0.10 in the 15-period games), while unsuccessful groups have a Gini coefficient of 0.27 (0.25 in 15-period games), more than twice as high. Hence, successful and unsuccessful groups already differ after the first round of the public good game.

This observation motivates us to study the extent of path dependency *within* groups. Table 5 presents evidence in this respect. It shows the correlation between wealth in period 10 (period 15) and the wealth or the Gini coefficient in previous periods. Path dependence is evident. Early period wealth is strongly correlated with late period wealth. The Spearman correlation coefficient between period 2 and 10 income is 0.48*** in NOPUNISH and even 0.82*** in PUNISH. Maybe more interestingly, also early period inequality is highly detrimental to final wealth, but only in PUNISH. The correlation between inequality in period 2 and wealth in period 10 is −0.47***, which is substantial. In NOPUNISH, by contrast, early period inequality (in periods 2-4) is not negatively correlated with wealth. In this treatment the negative correlation appears only from period 5 onwards. The fact that there is no negative correlation between period 7,8,9 inequality and period 10 wealth in PUNISH is due to the fact that by then several groups have zero wealth and inequality. Dropping these groups restores the negative correlation. Comparing 15 period wealth and early period inequality (bottom panel of Table 5) shows similar patterns.

	Wealth in period ...							
	9	8	7	6	5	4	3	2
NOPUNISH	0.99***	0.98***	0.98***	0.93***	0.85***	0.76***	0.75***	0.48***
PUNISH	0.99***	0.99***	0.98***	0.97***	0.93***	0.90***	0.89***	0.82***
	Gini coefficient in period ...							
	9	8	7	6	5	4	3	2
NOPUNISH	−0.29***	−0.16*	−0.35***	−0.46***	−0.34***	−0.03	0.03	0.11
PUNISH	0.03	0.11	−0.06	−0.43***	−0.17*	−0.32***	−0.50***	−0.47***
	Gini coefficient in period ... (15 period games)							
	14	13	12	11 ...	5	4	3	2
NOPUNISH	−0.12	−0.13	−0.12	−0.16	−0.19	0.07	−0.02	0.18
PUNISH	0.31*	−0.50***	−0.41***	−0.53***	−0.56***	−0.50***	−0.55***	−0.46***

Table 5: Correlation of wealth in period 10 with wealth (Gini) in period 2,...,9. Bottom panel: correlation of wealth in period 15 with Gini in periods 11-14 and 2-5, respectively.

One question about these findings is whether they just reflect a stable distribution of "contribution types" or whether there is something more fundamental to it in the sense that *the same* people are more likely to end up with a much lower wealth if initial inequality is high. The fact that the negative correlation between early period inequality and wealth is only observed in PUNISH, suggests that this is not just a mechanical effect of having different distributions of "contribution types" across groups.

Additional support for this view can be derived from our post-experimental questionnaire. Table 16 in Online Appendix F shows the average amount in Euros that participants decide to donate to Medics without Borders at the end of the experiment. We find that participants from above median groups do *not* contribute more on average than those from groups with below median wealth. This is despite the fact that participants from groups with above median wealth earn 178 tokens on average in period 10 (189 in treatment PUNISH), while those from groups with below median wealth earn only 56 tokens (23 tokens) on average in period 10. This suggests that participants in groups with above median wealth are *not* per se more altruistic than others. We also elicited 14 other personality characteristics as well as a measure of risk aversion in the questionnaire. Tables 17-18 in Online Appendix F show that none of them is able to explain the variation in wealth or inequality that we observe.

Result 5: Relation between Growth and Inequality. 1. *In PUNISH wealth and Gini coefficient are positively correlated for poor groups (below median wealth) and negatively correlated for rich groups (above median wealth) in the 10-period games. These correlations are weak and not statistically different from zero in NOPUNISH.*

2. *(a) Early period wealth is positively correlated with eventual wealth in all treatments. (b) Early period inequality is negatively correlated with final wealth only in PUNISH.*

Taken together both the findings on across- as well as within- group correlation point to a detrimental role of inequality if there are punishment possibilities. One of the reasons, hence, why punishment is not as effective in this setting seems to be that people react strongly to inequality. The findings are also indicative as to why we observe such substantial variation across groups both in terms of wealth and inequality. Groups in which initial behavior leads to high inequality seem to get locked into a path of punishment and counter-punishment (interpreted as experimental equivalents of conflict) that eventually leads to a destruction of all wealth. We will see more evidence of such behaviour in Section 5.3. In fact there is a strand of literature, where economic historians point out the importance of institutional lock in with institutions broadly understood as both formal constraints, such as rules and laws and informal constraints, such as norms of behaviour, conventions or codes of conduct (North, 1994). Our study provides an example of how a society (group) can get locked into dysfunctional behavioural norms.

Before we conclude this section, let us point to potential links to other literatures. The sign of the relation between inequality and growth as well as the causal link between the two has been at the center of a debate in macroeconomics and development (see e.g. Barro, 2000; Forbes, 2000; Persson and Tabellini, 1991, among many others). Most of these authors find either a negative relation or no significant relation at all. In our context the relation depends on the wealth of the group. For very poor groups in our data inequality and wealth are positively related, while they are negatively related for richer groups. This is reminiscent of the famous Kuznets curve (Kuznets, 1955) which claims an inverse U-shaped relationship between growth and inequality. The connection between our setting and the fate of countries is too loose to draw any conclusions. Our results suggest, however, that there may be interesting links, between the level of social capital (cooperation, trust), the level of inequality and the level of growth of societies.[11]

5.2 Eliminating inequality and growth possibilities

In this subsection we further tease apart the importance of endogenous inequality and growth by studying treatments were we shut down one of these channels exogenously. In treatments NOPUNISH-NOGROWTH and PUNISH-NOGROWTH we artificially eliminate growth by re-normalizing all endowments at the beginning of a period s.t. they sum to 80. More precisely, while in treatment NOPUNISH the following relation holds for endowments $N_i^{t+1} = N_i^t - a_i^t + \frac{1.5}{4}\sum_{j=1,\dots,4} a_j^t$, we now re-normalize

$$N_i^{t+1} = \frac{80 * \left(N_i^t - a_i^t + \frac{1.5}{4}\sum_{j=1,\dots,4} a_j^t\right)}{\sum_{j=1,\dots,4}\left(N_j^t - a_j^t + \frac{1.5}{4}\sum_{j=1,\dots,4} a_j^t\right)}.$$

This means that each participant's endowment ranges between 0 and 80 in each round and the sum of endowments in a group equals 80 in each period. This 10 period game hence, can be viewed as a public good game played in period 10, where the endowment each participant

[11] Knack and Keefer (1997) and Zak and Knack (2001) have already identified a relationship between trust and growth across different countries. Our setting and results point to a potential mechanism behind this relationship.

has is determined by the game played in periods 1-9. The unique SPNE is zero contributions in each period as in our main treatments. Note that this structure, while derived from our treatment NOPUNISH via a single change creates a situation where everyone contributing no longer Pareto dominates nobody contributing in periods 1-9. Consequently, participants may have different motives for contributing in these treatments compared to our main treatments and the results should be interpreted with this in mind. We do the analogous normalization for the punishment version (PUNISH-NOGROWTH). We had 116 participants (29 groups) in treatment NOPUNISH-NOGROWTH and 92 participants (23 groups) in treatment PUNISH-NOGROWTH.

Table 13 (column (1)) in Appendix D shows the results of regression where we regress normalized contributions on treatment dummies. The baseline is NOPUNISH. To ensure a fair comparison, we also normalize contributions in NOPUNISH and PUNISH. Normalized contributions in these treatments are computed by multiplying the share of actual endowment contributed with normalized endowments. Maybe unsurprisingly, given the absence of even social incentives to contribute across periods 1-9 of the games without growth, contributions are substantially lower and close to zero.

In treatments NOPUNISH-NOINEQUALITY and PUNISH-NOINEQUALITY we artificially eliminate inequality. In these treatments we redistribute all earnings at the beginning of each period, s.t. endowments are equal for all players. Hence each player receives

$$N_i^{t+1} = \frac{\sum_{i=1,\dots,4}\left(N_i^t - a_i^t + \frac{1.5}{4}\sum_{j=1,\dots,4} a_j^t\right)}{4}$$

at the beginning of period $t+1$. To reduce incentives to contribute, we implemented an additional change in this treatment. In particular we paid a randomly drawn period rather than the last period.[12] Of course paying a randomly selected period could affect behaviour *per se*, a possibility we cannot rule out. To this extent, the results from this treatment should be read as suggestive and as complementing the evidence from Sections 5.1 and 5.3 rather than yielding definite conclusions on the role of inequality. We had 96 participants (24 groups) in NOPUNISH-NOINEQUALITY and 56 participants (14 groups) in PUNISH-NOINEQUALITY.

The second column in Table 13 (Appendix D) shows that inequality leads to lower contributions in both the treatments with and without punishment. Eliminating inequality increases mean contributions by 26 tokens without punishment and by 42 tokens with punishment. Furthermore, if inequality is eliminated then mean contributions are ≈ 24 tokens higher with punishment (PUNISH-NOINEQUALITY) than without (NOPUNISH-NOINEQUALITY), a difference that is statistically significant ($\chi^2 < 0.001$). It seems that endogenous inequality in endowments with the associated inequality in the power to punish undermines the effectiveness of punishment. Since we exogenously remove inequality in the treatments discussed here, the results from this section support a causal interpretation of the effect of inequality on wealth.

5.3 Anatomy of Punishment

Both sections 5.1 and 5.2 have pointed to a negative role of inequality for contributions and wealth particularly in treatment PUNISH. To understand why this is the case it is helpful to

[12]Paying the last period in this treatment (as we did in the other treatments) would mean that there are incentives to contribute across all periods, making this setting difficult to compare to our other treatments. Paying a randomly selected period reduces this incentive. In particular, with a randomly paid period there are no incentives to contribute in the last period. In the 9-th period there is a tradeoff between increasing the endowment available in period 10 by contributing and increasing the 9th period payoff by no contributing. This tradeoff is resolved towards not contributing for periods 9,8,7 and 6. In earlier periods, however, there are incentives to contribute in order to increase the endowment available in future rounds.

study punishment patterns more closely.

We first compare above and below median groups in treatment PUNISH to understand which patterns of punishment lead to low wealth in this treatment and are hence crucial for the ineffectiveness of punishment in this setting. Figure 4 reveals an interesting pattern in this regard. It shows the amounts of tokens participants use to punish over time. In groups with above median wealth the absolute amounts used to punish tend to remain stable or increase over time (OLS coefficient: 1.402^{*} (10 periods); 0.098 (15 periods)),[13] while they tend to decrease in below median groups (OLS coefficient: -0.830^{***} (10 periods); -0.332^{***} (15 periods)). Particularly striking is the fact that, in terms of amounts, the major difference between above and below median groups seems to lie in how much they punish in the first two periods of the game (two-sided ranksum test $p < 0.0001$). In successful groups, there is also an interesting peak in punishment one period before the game ends. This suggests that some participants may tolerate some degrees of free-riding while they wait to punish others harshly at the end of the game. This seems intuitive because of the detrimental effect that punishment can have on growth. While in below median groups most punishment happens in the beginning of the game, above median groups punish at the end.[14]

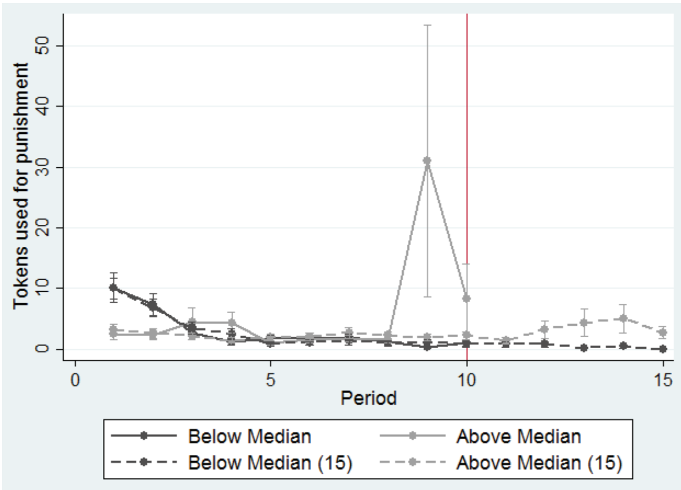


Figure 4: Average Amount of Tokens used to punish over time by groups with above median wealth (dashed line) and groups with below median wealth (solid line).

We next ask under which conditions punishment is "*pro-social*" and when it is "*anti-social*". Herrmann et al. (2008) have found that anti-social punishment strongly undermines successful cooperation in the standard setting, whereas pro-social punishment fosters cooperation. Punishment by player i is *pro-social* if i punishes a player who has contributed a lower share of his endowment to the public good than i herself. Punishment by player i is *antisocial* if i punishes a player who has contributed a higher share of his endowment to the public good than i herself.

Table 6 shows that there is more pro-social punishment than anti-social punishment (one-sided ranksum test, $p < 0.0001$). Anti-social punishment is higher if inequality is high (pro-social: Spearman $\rho = 0.0155$ ($\rho = 0.1231^{***}$, 15 period games); anti-social: $\rho = 0.0777^{**}$ ($\rho = 0.1515^{***}$, 15 period games)). Anti-social punishment is not significantly related to wealth ($\rho = -0.0109$; $\rho = 0.1184$ in 15 period games), but there is more pro-social punishment in high-wealth compared to low-wealth groups ($\rho = 0.1371^{***}$; $\rho = 0.2302^{***}$ in 15 period games).

[13]We regress total amount spent on (intended) punishment on a constant and a variable period ranging from 1,...,10 (15, respectively).

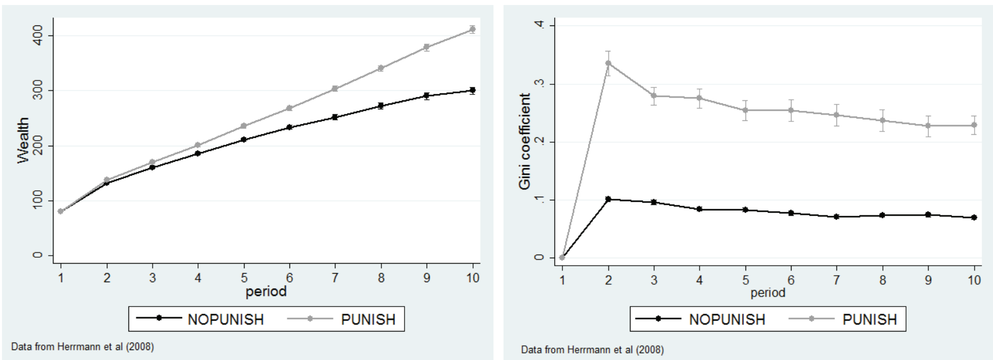
[14]Relatedly, Fudenberg and Pathak (2010) have shown that punishment can sustain cooperation even if it is only observed at the end of a session.

	10 period games	
	Pro-social Punishment	Antisocial Punishment
all	1.91 (18.73)	0.65 (4.07)
Gini> 0.04	1.14 (3.24)	1.52 (12.03)
Gini< 0.04	2.68 (26.22)	0.45 (2.10)
wealth> 80	3.10 (26.84)	0.62 (5.25)
wealth< 80	0.44 (1.58)	0.46 (1.87)
	15 period games	
	Pro-social Punishment	Antisocial Punishment
all	1.24 (5.01)	0.65 (2.83)
Gini> 0.08	1.25 (4.23)	0.84 (3.55)
Gini< 0.08	1.23 (5.55)	0.50 (2.09)
wealth> 158	2.25 (8.26)	0.62 (3.84)
wealth< 158	0.77 (2.14)	0.66 (2.22)

Table 6: Mean (SD) of amount of tokens of pro-social and anti-social punishment.

5.4 Wealth and Inequality in the Standard Setting

Before we conclude, we have a brief look at data from the standard setting. We focus on wealth and inequality, as those are not commonly reported measures in the standard setting. In particular, we use data from experiments Herrmann et al. (2008) conducted in Bonn. Bonn and Maastricht are at only 120km driving distance and a large share of students at both Universities come from the lower middle Rhine and upper lower Rhine area (between Koblenz and Duesseldorf) in Germany. This should make subject pools approximately comparable. Just as us Herrmann et al. (2008) conducted 10 period games with an endowment of 20 tokens (in their case exogenously given each period). They also conducted NOPUNISH and PUNISH treatments with the 3:1 technology that we employ. The return on contributing in the one-shot game was 0.375 in our experiments and at 0.4 slightly higher in Herrmann et al. (2008).



(a) Wealth (b) Inequality

Figure 5: Wealth and Inequality in data from Herrmann et al. (2008)'s Bonn experiments. Error bars show standard errors of the mean.

Figure 5 shows wealth and inequality (Gini coefficient) over time in Herrmann et al. (2008)'s experiments. We computed wealth by adding income of all group members in each period and subtracting 80 in each period 2,...,10. If all group members contribute zero in each period (and do not punish) wealth thus computed is 80 in each period, just as in our setting. If all group members contribute their full endowment in each period (and do not punish) maximal wealth in period 10 is 3075 in our setting and 1280 in Herrmann et al. (2008). Figure 5 shows that both wealth and inequality are substantially higher under the punishment condition in Herrmann et al. (2008)'s experiments (t-test period 10 sample means: $p < 0.01$). The effect of punishment, hence, differs from what we observe in the dynamic setting, where punishment, if at all, decreases wealth and inequality. Note also, that we do not observe the cyclical pattern in inequality identified in the PUNISH treatments in our setting.

In summary, section 5.1 demonstrates a negative correlation between early period inequality and wealth in period 10 in treatment PUNISH. The negative impact of inequality on contributions and wealth is confirmed with some caveats in Section 5.2 where we find that artificially removing inequality increases contributions and wealth both with and without punishment, but with a particularly strong effect in the punishment treatments. Section 5.3 provides some insights into why inequality is so detrimental in the punishment condition. There are two main effects: (i) in groups where inequality is high (above median) there is more anti-social than pro-social punishment, implying that shirkers punish contributors more than vice versa and (ii) much of this punishment happens in early periods in unsuccessful groups implying that resources are taken away exponentially. All these channels contribute to the negative impact of punishment possibilities on wealth creation in the dynamic setting. The absence of inequality in endowments in the standard setting seems to make wealth inequality less salient. This can potentially explain why punishment does not trigger the adverse affects identified in the dynamic setting in the standard setting (Section 5.4).

6 Conclusions

We studied public good games with dynamic interdependencies, where each agent's wealth at the end of a period serves as her endowment in the following period. We found that contributions are increasing over time even in the absence of punishment possibilities. The possibility of punishment does not increase wealth. These results suggest that in settings with a strong dynamic component societies achieve cooperation via the incentives provided by the dynamic evolution of the public good and not so much via the threat of punishment. Across groups, inequality in early periods is strongly negatively correlated with wealth in later periods.

These results show that people are able to establish persistent cooperation in a setting that shares one key feature with many real-life interactions: past behaviour matters for future endowments. They also point to the limits of punishment in securing high contributions and wealth.

The results highlight the importance of incorporating the dynamic aspects, present in many real-life interactions, explicitly in experimental designs. In this paper we have done so within the context of cooperation and public good provision. We should emphasize, though, that we view this design as complementary to the standard design where contributions occur from stationary exogenous budgets. Both settings have natural applications outside the lab. Public goods are often created by providing effort, as e.g. in the case of volunteering. In these cases, clearly, endowments cannot easily increase over time. In many other settings, such as those discussed in the Introduction, provision of public goods creates wealth from which future public goods can be provided. The evolution of societies might be viewed through this lens. We have seen that the two settings can yield quite different conclusions. In the standard setting with stationary budgets and full consumption punishment has shown to be very effective at raising contributions and wealth (in the medium run). In the dynamic setting studied here, the effectiveness of punishment seems much more limited. It is not successful in raising wealth in the 10-period games and is even detrimental to wealth in the games with longer horizon.

Future research should aim at getting a deeper understanding of the mechanisms behind these differences and adapt this setting to other contexts where dynamic interdependencies are likely to play an important role. Studying intermediate settings, with some but not full consumption and only partially endogenous endowments should also be of interest for future research.

References

Adler, P. (1985). *Wheeling and Dealing - An Ethnography of Upper-Level Drug Dealing and Smuggling Community*. Columbia University Press.

Andreoni, J. (1995). Cooperation in public goods experiments: Kindness or confusion? *The American Economic Review 85(4)*, 891–904.

Andreoni, J., W. Harbaugh, and L. Vesterlund (2003). The carrot or the stick: rewards, punishment and cooperation. *The American Economic Review 93(3)*, 893–902.

Barro, R. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth 5*, 5–32.

Battaglini, M., S. Nunnari, and T. Palfrey (2014). The dynamic free rider problem: A laboratory study. mimeo.

Battigalli, P. and M. Siniscalchi (1999). Hierarchies of conditional beliefs and interactive epistemology in dynamic games. *Journal of Economic Theory 88*, 188–230.

Battigalli, P. and M. Siniscalchi (2002). Strong belief and forward induction reasoning. *Journal of Economic Theory 106*, 356–391.

Becker, G. (1981). *A treatise on the family*. Cambridge.: Harvard University Press.

Bergstrom, T., L. Blume, and H. Varian (1986). On the private provision of public goods. *Journal of Public Economics 29*, 25–49.

Botelho, A., A. Dinar, L. C. Pinto, and A. Rapoport (2014). Time and uncertainty in resource dilemmas: equilibrium solutions and experimental results. *Experimental Economics 17*, 649–672.

Cadigan, J., P. Wayland, P. Schmitt, and K. Swope (2011). An experimental dynamic public goods game with carryover. *Journal of Economic Behavior and Organization 80*, 523–531.

Cason, T. and L. Gangadharan (2015). Promoting cooperation in nonlinear social dilemmas through peer punishment. *Experimental Economics 18*, 66–88.

Chan, K., S. Mestelman, R. Moir, and R. Muller (1996). The voluntary provision of public goods under varying income distributions. *The Canadian Journal of Economics 29(1)*, 54–69.

Chan, K. S., S. Mestelman, R. Moir, and R. Muller (1999). Heterogeneity and the voluntary provision of public goods. *Experimental Economics 2(1)*, 5–30.

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

Cozzi, G. (1999). R+d cooperation and growth. *Journal of Economic Theory 86(1)*, 17–49.

Croson, R. and M. Marks (1998). Alternative rebate rules in the provision of a threshold public good: An experimental investigation. *Journal of Public Economics 67(2)*, 195–220.

Dawes, R., J. Orbell, R. Simmons, and A. J. van de Kragt (1986). Organizing groups for collective action. *The American Political Science Review 80(4)*, 1171–1185.

Deaton, A. (1997). *Analysis of Household Surveys*. Baltimore, MD:: Johns Hopkins University Press.

Dreber, A., D. Rand, D. Fudenberg, and M. Nowak (2008). Winners don't punish. *Nature 452*, 348–351.

Duffy, J., J. Ochs, and L. Vesterlund (2007). Giving little by litte: Dynamic voluntary contribution games. *Journal of Public Economics 91(9)*, 1708–1730.

Egas, M. and A. Riedl (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B - Biological Sciences 275*, 871–878.

Fehr, E. and S. Gächter (2000). Cooperation and punishment in public goods experiments. *American Economic Review 90*, 341–369.

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

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

Fischbacher, U. and S. Gächter (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *The American Economic Review 100(1)*, 541–556.

Forbes, K. (2000). A reassessment of the relationship between inequality and growth. *American Economic Review 90(4)*, 869–997.

Fudenberg, D. and P. Pathak (2010). Unobserved punishment supports cooperation. *Journal of Public Economics 94(1)*, 78–86.

Gächter, S., E. Renner, and M. Sefton (2008). The long-run benefits of punishment. *Science 322*, 1510.

Goeree, J. K., C. Holt, and S. K. Laury (2002). Private costs and public benefits: unraveling the effects of altruism and noisy behavior. *Journal of Public Economics 83*, 255–276.

Green, E. and R. Porter (1984). Noncooperative collusion under imperfect price information. *Econometrica 52(1)*, 87–100.

Greiner, B., A. Ockenfels, and P. Werner (2011). The dynamic interplay of inequality and trust - an experimental analysis. *Journal of Economic Behavior and Organization 81(2)*, 355–365.

Grossman, G. and C. Shapiro (1986). Research joint ventures: An anti-trust analysis. *Journal of Law, Economics and Organization 2(2)*.

Guererk, O., B. Rockenbach, and I. Wolff (2017). The effects of punishment in dynamic public-good games. working paper 106 Thurgau Institute of Economics.

Hamilton, B. H., J. A. Nickersen, and H. Owan (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy 111*, 465–497.

Herr, A., R. Gardner, and J..M.Walker (1997). An experimental study of time-independent and time-dependent externalities in the commons. *Games and Economic Behavior 19*, 77–96.

Herrmann, B., C. Thöni, and S. Gächter (2008). Antisocial punishment across societies. *Science 319*, 1362–1367.

Huck, T. (2006). Dynamisches public good experiment. Master's thesis, Universitaet Erfurt.

Isaac, R. and J. Walker (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *Quarterly Journal of Economics 103(1)*, 179–199.

Isaac, R. M., J. M. Walker, and S. Thomas (1984). Divergent evidence on free-riding: An experimental examination of possible explanations. *Public Choice 43(2)*, 113–149.

Johnson, A. and T. Earle (1987). *The Evolution of Human Societies: From Foraging Group to Agrarian State*. Stanford University Press.

Josephson, M. (1962). *The Robber Barons*. Mariner Books.

Kamien, M. I., E. Muller, and I. Zang (1992). Research joint ventures and r+d cartels. *The American Economic Review 82(5)*, 1293–1306.

Kimbrough, E. and A. Vostroknutov (2015). The social and ecological determinants of common pool resource sustainability. *Journal of Environmental Economics and Management 72*, 38–53.

Knack, S. and P. Keefer (1997). Does social capital have an economic payoff?: a cross country investigation. *The Quarterly Journal of Economics 112(4)*, 1251–1288.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory 27*, 245–252.

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review 45(1)*, 1–28.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel and A. E. Roth (Eds.), *Handbook of Experimental Economics*. Princeton University Press.

Mailath, G. and L. Samuelson (2005). *Repeated Games and Reputation*. Oxford University Press.

McKelvey, R. and T. Palfrey (1992). An experimental study of the centipede game. *Econometrica 60(4)*, 803–836.

Milinski, M., D. Semmann, and H.-J. Krambeck (2002). Reputation helps solve the 'tragedy of the commons'. *Nature 415*, 424–426.

Nishi, A., H. Shirado, D. G. Rand, and N. Christakis (2015). Inequality and visibility of wealth in experimental social networks. *Nature 526*.

North, D. (1994). Performance through time. *The American Economic Review 84(3)*, 359–368.

Noussair, C. and C. Soo (2008). Voluntary contributions to a dynamic public good: Experimental evidence. *Economics Letters 98*, 71–77.

Ostrom, E. (1990). *Governing the Commons - The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

Ostrom, E., R. Gardner, and J. Walker (1994). *Rules, Games and Common Pool Ressources*. Ann Arbor: University of Michigan Press.

Ostrom, E., J. Walker, and R. Gardner (1992). Covenants without a sword: Self-governance is possible. *The American Political Science Review 86*, 104–115.

Persson, T. and G. Tabellini (1991). Is inequality harmful for growth? theory and evidence. *American Economic Review 48*, 600–621.

Plott, C. R. and V. Smith (Eds.) (2009). *Handbook of Experimental Economics Results*. North-Holland.

Reuben, E. and A. Riedl (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior 77*, 122–137.

Sadrieh, A. and H. Verbon (2006). Inequality, cooperation and growth: an experimental study. *European Economic Review 50(5)*, 1197–1222.

Tirole, J. (1986). Hierarchies and bureaucracies: On the role of collusion in organizations. *Journal of Law, Economics and Organization 2(2)*, 181.

van Dijk, F., J. Sonnemans, and F. van Winden (2002). Social ties in a public good experiment. *Journal of Public Economics 85*, 275–299.

Zak, P. J. and S. Knack (2001). Trust and growth. *Economic Journal 111*, 295–321.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics 6*, 299–303.

Online Appendix for "Growth and Inequality in Public Good Provision"

S. Gächter; F. Mengel; E. Tsakas and A. Vostroknutov

Contents

A Screenshots

	Tokens at the beginning of this period	Tokens placed in group account	Share of group account return	Tokens in the end of the period
You	9	5.00	4.87	9
Other1	11	2.00	4.87	14
Other2	12	6.00	4.87	11
Other3	1	0.00	4.87	6

Figure A.1: Information the participants see after each period in NOPUNISH treatment. Fictitious participant "You" contributed 5 of his 9 tokens and received a share of 5 (4.87 rounded up) from the group account resulting in $9 - 5 + 5 = 9$ tokens before punishment.

	Tokens at the beginning of this period	Tokens placed in group account	Share of group account return	Total tokens before subtraction	Number of tokens to subtract
You	9	5.00	4.87	9	
Other1	11	2.00	4.87	14	
Other2	12	6.00	4.87	11	
Other3	1	0.00	4.87	6	

Figure A.2: The screen shot of the punishment stage. The assignment of "Other" categories were randomized in each period.

	Tokens at the beginning of this period	Tokens placed in group account	Share of group account return	Total tokens before subtraction	You subtracted	You got subtracted by others	Sum of subtractions by others	Tokens in the end of the period
You	9	5.00	4.87	9			4.00	3
Other1	11	2.00	4.87	14	1.00	1.00	4.00	9
Other2	12	6.00	4.87	11	2.00	0.00	4.00	6
Other3	1	0.00	4.87	6	3.00	3.00	6.00	0

Figure A.3: Information available to the participants after the punishment phase. The fictitious participant "You" punished "Other1" by 1, "Other 2" by 2 and was punished 1 (*3) by "Other 1" resulting in 9-3-3=3 tokens at the end of the period. Plans regarding "Other 3" were not executed (indicated by a pop up window not shown here), because player 3 was already set to zero.

B Questionnaire and Experimental Instructions

B.1 Instructions NOPUNISH Treatment

General information

You are about to participate in a decision making experiment. If you follow the instructions carefully, you can earn a considerable amount of money depending on your decisions and the decisions of the other participants. Your earnings will be paid to you in cash at the end of the experiment

This set of instructions is for your private use only. During the experiment you are not allowed to communicate with anybody. In case of questions, please raise your hand. Then we will come to your seat and answer your questions. Any violation of this rule excludes you immediately from the experiment and all payments. The funds for conducting this experiment were provided by the Marie Curie Reintegration Grant from the EU.

Throughout the experiment you will make decisions about amounts of tokens. At the end of the experiment all tokens you have will be converted into Euros at the exchange rate 0.05 Euro for 1 token and paid you in cash in addition to the show-up fee of 2 Euros.

During the experiment all your decisions will be treated confidentially. This means that none of the other participants will know which decisions you made.

Experimental Instructions

The experiment will consist of 10 decision making periods. At the beginning of the experiment, you will be matched with 3 other people in this room. Therefore, there are 4 people, including yourself, participating in your group. You will be matched with the same people during the entire experiment. None of the participants knows who is in which group.

Before the first period you, and each other person in your group, will be given the endowment of 20 tokens.

At the beginning of the first period you will be asked to allocate your endowment between a private account and a group account.

The tokens that you place in the private account have a return of 1 at the end of the first period. This means that at the end of the first period your private account will contain exactly the amount of tokens you put into the private account at the beginning of the period. Nobody except yourself benefits from your private account.

The tokens that you place in the group account are summed together with the tokens that the other three members of your group place in the group account. The tokens in the group account have a return of 1.5. Every member of the group benefits equally from the group account. Specifically, the total amount of tokens placed in the group account by all group members is multiplied by 1.5 and then is equally divided among the four group members. Hence, your share of the group account at the end of the first period is

1.5 * (sum of tokens in the group account) / 4

Your endowment at the beginning of the second period will be equal to the amount of tokens contained in your private account at the end of the first period plus your share of the group account at the end of the first period.

At the beginning of the second period you will be again asked to allocate the endowment that you have at the beginning of the second period between a private account and a group account. Both the private and the group account work in exactly the same manner as in the first period, namely, they have the same returns.

The structure of the experiment at all subsequent periods is identical: your endowment at the beginning of each period is equal to the amount of tokens in your private account at the end of the previous period plus your share of the group account at the end of the previous period.

At the end of each period, you will be informed about

- The endowment all four group members had at the beginning of the period

- How much each group member allocated to the group account and to their respective private accounts.
- Your share of the group account (remember it is the same for all group members).

All other participants will receive exactly the same information.

Your total income in the end of the experiment is equal to the amount of tokens in your private account and your share of the group account at the end of period 10. At the end of the experiment there will be a short questionnaire for you to fill in.

B.2 Instructions PUNISH Treatment

General information

You are about to participate in a decision making experiment. If you follow the instructions carefully, you can earn a considerable amount of money depending on your decisions and the decisions of the other participants. Your earnings will be paid to you in cash at the end of the experiment

This set of instructions is for your private use only. During the experiment you are not allowed to communicate with anybody. In case of questions, please raise your hand. Then we will come to your seat and answer your questions. Any violation of this rule excludes you immediately from the experiment and all payments. The funds for conducting this experiment were provided by the Marie Curie Reintegration Grant from the EU.

Throughout the experiment you will make decisions about amounts of tokens. At the end of the experiment all tokens you have will be converted into Euros at the exchange rate 0.05 Euro for 1 token and paid you in cash in addition to the show-up fee of 2 Euros.

During the experiment all your decisions will be treated confidentially. This means that none of the other participants will know which decisions you made.

Experimental Instructions

The experiment will consist of 10 decision making periods. Each period consists of two stages. At the beginning of the experiment, you will be randomly matched with 3 other people in this room. Therefore, there are 4 people, including yourself, participating in your group. You will be matched with the same people during the entire experiment. None of the participants knows who is in which group.

Before the first period you, and each other person in your group, will be given the endowment of 20 tokens.

At the first stage of the first period you will be asked to allocate your endowment between a private account and a group account.

The tokens that you place in the private account have a return of 1 at the end of the first stage. This means that at the end of the first stage your private account will contain exactly the amount of tokens you put into the private account at the beginning of the first stage. Nobody except yourself benefits from your private account.

The tokens that you place in the group account are summed together with the tokens that the other three members of your group place in the group account. The tokens in the group account have a return of 1.5. Every member of the group benefits equally from the tokens in the group account. Specifically, the total amount of tokens placed in the group account by all group members is multiplied by 1.5 and then is equally divided among the four group members. Hence, your share of the group account at the end of the first stage of the first period is

1.5 * (sum of tokens in the group account) / 4

In the second stage of the first period you will be asked to react to the decisions made during the first stage of the first period. At this point, you will already know the decisions taken by each group member at the first stage. You will decide whether you want to subtract tokens from any other group member or not. The members that you decide to subtract tokens from will lose the amount of tokens you chose. Subtracting tokens from someone else is costly for you too. The following table illustrates the relation between your cost in tokens and the amount of tokens that are taken away from the member of your group:

Tokens subtracted	Cost for you
3	1
6	2
9	3
...	...
$3y$	y

You may subtract different amounts of tokens from different group members. Other group members will be able to subtract tokens from you as well. You lose the sum of tokens that other three group members decided to subtract from you. Any group member including you can only lose maximum the amount of tokens he or she has.

At the beginning of the second period your endowment will be equal to the amount of tokens contained in your private account at the end of the first stage of the first period plus your share of the group account at the end of the first stage of the first period, minus your cost for subtracting others' tokens and minus the amount of tokens subtracted from you by other members.

At the first stage of the second period you will be again asked to allocate the endowment that you have at the beginning of the second period between a private account and a group account. Both the private and the group account work in exactly the same manner as in the first period, namely, they have the same returns. At the second stage of the second period you will be asked to react to the decisions made during the first stage of the second period in exactly the same manner as in the first period.

The structure of the experiment at all subsequent periods is identical: your endowment at the beginning of each period is equal to the amount of tokens in your private account at the end of the first stage of previous period, plus your share of the group account at the end of the first stage of the previous period, minus your cost from subtracting other members' tokens at the second stage of the previous period, minus the amount of tokens subtracted from you by other members at the second stage of the previous period.

At the end of each period, you will be informed about

- The endowment all four group members had at the beginning of the period
- How much each group member allocated to the group account and to their respective private accounts
- Your share of the group account (remember it is the same for all group members)
- How many tokens each member subtracted from you.

All other participants will receive exactly the same instructions.

Your total income in the end of the experiment is equal to the amount of tokens left after last subtraction in your private account and your share of the group account at the end of period 10. At the end of the experiment there will be a short questionnaire for you to fill in.

B.3 Questionnaire

The following questions were asked after both PUNISH and NOPUNISH treatments.

- What is your gender?
- What is your nationality?
- What is your year of birth?
- What is your field of studies?

- For how many years have you been studying at university?

Suppose you have a hypothetical choice between a bet and a sure outcome. What would you choose in the following cases:

- €10 Euro or 100 Euro with 50% chance and €0 Euro with 50% chance
- €20 Euro or 100 Euro with 50% chance and €0 Euro with 50% chance
- €30 Euro or 100 Euro with 50% chance and €0 Euro with 50% chance
- €40 Euro or 100 Euro with 50% chance and €0 Euro with 50% chance
- €50 Euro or 100 Euro with 50% chance and €0 Euro with 50% chance
- €60 Euro or 100 Euro with 50% chance and €0 Euro with 50% chance
- €70 Euro or 100 Euro with 50% chance and €0 Euro with 50% chance

Personality questions: indicate how strongly you agree with the following statements (1 means disagree strongly, 7 agree strongly).

- I am a quick thinker
- I get easily offended
- I am very satisfied with myself
- I am very dependent on others
- Generally speaking, I am happy
- Work plays a very important role in my life
- Family plays a very important role in my life
- Friends play a very important role in my life
- Religion plays a very important role in my life
- Politics plays a very important role in my life
- Generally, most people can be trusted
- In the long run, hard work brings a better life
- The government should take responsibility that people are better provided for
- Incomes should be made more equal

In addition the participants were asked if they would be willing to donate some of their earnings to Doctors without Borders.

C Proofs of Section 3

In this section we establish the theoretical results discussed informally in Section 3. We start with some notation, then prove our first results on the structure of NE and SPE in our public good games with growth. Finally, we expand our model by introducing behavioral types which allow us to model reputation effects, and we provide a result on the structure of sequential equilibria in the induced incomplete information game.

C.1 Notation and preliminaries

Players and histories. Let I be the set of four players. Each player receives a random index $i \in \{1, \dots, 4\}$ at the beginning of the game. Hereinafter we identify each player with the respective index. Furthermore, let H denote the set of non-terminal histories and Z denote the set of terminal histories. There are two types of non-terminal histories, contribution and punishment histories, denoted by H^c and H^p respectively. The root of the game is a contribution history, i.e., $h_1 \in H^c$. Games without punishment contain only contribution histories. On the other hand games with punishment contain both contribution and punishment histories. In the latter case, the two types of histories occur in an alternating order, i.e., the direct predecessor of each $h \in H^c \setminus \{h_1\}$ belongs to H^p, and vice versa, the direct predecessor of each $h \in H^p$ belongs to H^c. Moreover, the final non-terminal history is a punishment history, i.e., the direct predecessor of each $z \in Z$ belongs to H^p.

Paths of play. A path is a sequence of histories beginning with the root of the game h_1, ending at a terminal history $z \in Z$, and containing a unique immediate successor for each non-terminal history, i.e., it is the collection of z's predecessors. Hence, a path is uniquely determined by the respective terminal history. We define the length of a path to be the number of (terminal and non-terminal) histories. Obviously, in our public good game without punishment each path is of length $T+1$, where T is the number of periods, i.e., a path contains a single (contribution) history for each period. In our public good game with punishment each path is of length $2T+1$, where T is again the number of periods, i.e., a path contains two histories for each period, a contribution history and a subsequent punishment history. In both cases, let H_t denote the non-terminal histories at stage t. Thus, in the game without punishment, a path is a sequence $(h_1, h_2, \dots, h_T, z)$ such that $h_t \in H_t$. On the other hand in the game with punishment a path is a sequence $(h_1^c, h_1^p, h_2^c, h_2^p, \dots, h_T^c, h_T^p, z)$ such that $h_t^c \in H_t^c := H_t \cap H^c$ and $h_t^p \in H_t^p := H_t \cap H^p$.

Strategies. Let A_i^h be the finite set of actions that player i has at $h \in H$. If h is a contribution history,

$$A_i^h := C_i^h := \{0, \dots, N_i^h\},$$

where N_i^h denotes the number of tokens in i's private account upon reaching the contribution history h. Obviously, A_i^h depends on the amount of tokens that player i has accumulated in her private account so far. This is for instance why our public good game without punishment is not a repeated game, as opposed to the standard case where $N_i^h = 20$ for all $h \in H^c$. If, on the other hand, h is a punishment history,

$$A_i^h := P_i^h := \Big\{ (p_{i,j})_{j \neq i} \in \mathbb{N}^3 : \sum_{j \neq i} p_{i,j} \leq W_i^h \Big\}$$

where $3p_{i,j}$ is the number of tokens that i subtracts from j's private account, and W_i^h is the number of tokens in i's private account upon reaching the punishment history h.[1] As usual, let $A_i := \prod_{h \in H} A_i^h$ denote the set of i's strategies and $A := \prod_{i \in I} A_i$ denote the set of strategy profiles. For an arbitrary $a \in A$, let $c_i^h(a) := a_i^h$ be i's action at the contribution history $h \in H^c$. Likewise, let $p_i^h(a) := a_i^h$ denote i's action at the punishment history $h \in H^p$.

An arbitrary strategy profile $a \in A$ induces a unique path $H(a)$. In a game without punishment, let $(c_i^1(a), \dots, c_i^T(a))$ denote i's observed actions (contributions) along the path $H(a)$. Likewise, in a

[1] Obviously, player i's total cost from punishing cannot exceed the number of tokens in her private account at the corresponding history.

game with punishment let $(c_i^1(a), p_i^1(a), \ldots, c_i^T(a), p_i^T(a))$ denote i's observed actions (contributions and subsequent punishments) along the path $H(a)$.

Payoff functions. Let us begin with our *public good game without punishment*: Fix an arbitrary strategy profile $a \in A$, and take each player i's observed contributions $(c_i^1(a), \ldots, c_i^T(a))$ along the realized path $H(a)$. Then, for each $t \geq 1$, we inductively define

$$N_i^{t+1} := N_i^t - c_i^t(a) + \frac{r}{4} \sum_{j=1}^{4} c_j^t(a), \tag{1}$$

with $N_i^1 := N_i^{h_1} = 20$ and r denoting the returns of the public good. Then, we define player i's payoff function $u_i : A \to \mathbb{R}$ in by

$$u_i(a) = N_i^{T+1}. \tag{2}$$

Now, fix an arbitrary $h \in H$ and an arbitrary strategy profile $a \in A$. Then let $a' \in A$ be a strategy profile – possibly other than a – such that (i) $h \in H(a')$, and (ii) a' agrees with a at all histories weakly following h. Then, we define player i's payoff from a conditional on the history h by

$$u_i(a|h) = u_i(a'). \tag{3}$$

Now, let us switch our focus to our *public good game with punishment*: Consider an arbitrary strategy profile $a \in A$, and take each player i's observed contributions $(c_i^1(a), p_i^1(a), \ldots, c_i^T(a), p_i^T(a))$ along the realized path $H(a)$. Recall that at the beginning of the game the players are ordered from 1 to 4, i.e., each player has received a unique index $i \in \{1, \ldots, 4\}$. Now, let $\mathcal{J}_i$ be the collection of nonempty subsets $J \subseteq I$ such that (a) $i \notin J$ and (b) if $k \in J$ then $j \in J$ for all $j \in \{1, \ldots, k\} \setminus \{i\}$. Then, for each $t \geq 1$, we inductively define

$$W_i^t := N_i^t - c_i^t(a) + \sum_{j \in I} c_j^t(a) \tag{4}$$

$$N_i^{t+1} := \min_{J \in \mathcal{J}_i^t} \left\{ W_i^t - \sum_{j \in J} 3p_{j,i}^t(a) \right\} - \sum_{j \neq i} p_{i,j}^t(a) \tag{5}$$

with $N_i^1 := N_i^{h_1} = 20$. Then, we define i's payoff function $u_i^p : A \to \mathbb{R}$ by

$$u_i^p(a) = N_i^{T+1}. \tag{6}$$

Now, once again fix an arbitrary $h \in H$ and an arbitrary strategy profile $a \in A$. Then, similarly to the game without punishment, $a' \in A$ be a strategy profile such that (i) $h \in H(a')$, and (ii) a' agrees with a at all histories weakly following h. Then, define player i's payoff from a conditional on the history h by

$$u_i^p(a|h) = u_i^p(a'). \tag{7}$$

Finally, note that in all our cases, we assume $r = 1.5$.

C.2 Predicted behavior: Results and Proofs

In the game without punishment (resp., with punishment) we say that a strategy $a_i \in A_i$ is a a best response to $a_{-i} \in A_{-i}$, and we write $a_i \in BR_i(a_{-i})$, whenever $u_i(a_i, a_{-i}) \geq u_i(b_i, a_{-i})$ (resp., whenever $u_i^p(a_i, a_{-i}) \geq u_i^p(b_i, a_{-i})$) for all $b_i \in A_i$. The strategy profile a is a Nash equilibrium (NE) whenever $a_i \in BR_i(a_{-i})$ for every $i \in I$. Likewise, in the game without punishment (resp., with punishment) we say that a strategy $a_i \in A_i$ is a best response to $a_{-i} \in A_{-i}$ conditionally on h, and we write $a_i \in BR_i(a_{-i}|h)$, whenever $u_i(a_i, a_{-i}|h) \geq u_i(b_i, a_{-i}|h)$ (resp., $u_i^p(a_i, a_{-i}|h) \geq u_i^p(b_i, a_{-i}|h)$), for all $b_i \in A_i$. The strategy profile a is a subgame perfect equilibrium (SPE) whenever $a_i \in BR_i(a_{-i}|h)$ for every $i \in I$ and every $h \in H$. It is well-known that in games with observable actions, SPE are consistent with the backward induction procedure.

Proposition 3 *Consider the public good game (with growth as defined above) without punishment.*

(i) The unique SPE is such that every player contributes 0 at every history, i.e., if $a \in A$ is a SPE, then $c_i^h(a) = 0$ for every $i \in I$ and for all $h \in H$.

(ii) Every NE is such that every player contributes 0 at every history on the equilibrium path, i.e., if $a \in A$ is a NE, then $c_i^h(a) = 0$ for every $i \in I$ and for all $h \in H(a)$.

Proof. (i) This part of the proof follows the standard backward induction argument. Let $a \in A$ be an SPE. For an arbitrary $t \in \{1, \ldots, T\}$, suppose that $c_i^{h'}(a) = 0$ for all $i \in I$ and all $h' \in H_{t+1} \cup \cdots \cup H_T$. Of course, if $t = T$, then our previous assumption becomes trivially vacuous. Now for an arbitrary $h \in H_t$, it suffices to prove that $c_i^h(a) = 0$ for all $i \in I$. Assume that this is not the case, i.e., assume that there is some $i \in I$ such that $c_i^h(a) > 0$. Take another strategy $b_i \in A_i$ such that $c_i^{h''}(b_i, a_{-i}) = c_i^{h''}(a)$ at every $h'' \neq h$ moreover $c_i^h(b_i, a_{-i}) = 0$. This implies that $c_j^{h'}(a) = c_j^{h'}(b_i, a_{-i}) = 0$ for all $h' \in H_{t+1} \cup \cdots \cup H_T$, and therefore i's private account will contain at the end of the game the amount of tokens it will contain after the history h. Hence,

$$
\begin{aligned}
u_i(a_i, a_{-i}|h) &= N_i^h - c_i^h(a) + \frac{r}{4}\sum_{j=1}^{4} c_j^h(a) \\
&< N_i^h + \frac{r}{4}\sum_{j \neq i} c_j^h(a) \\
&= N_i^h - c_i^h(b_i, a_{-i}) + \frac{r}{4}\sum_{j=1}^{4} c_j^h(b_i, a_{-i}) \\
&= u_i(b_i, a_{-i}|h)
\end{aligned}
$$

thus implying that $a_i \notin BR_i(a_{-i}|h)$ and therefore a is not an SPE, which contradicts our hypothesis above. Hence, $c_i^h(a) = 0$ for all $i \in I$ and all $h \in H_t$, which completes the proof.

(ii) Let $a \in A$ be a NE, and recall that by $(c_i^1(a), \ldots, c_i^T(a))$ we denote i's observed actions along the equilibrium path $H(a)$. Now, suppose that there exists some $t \in \{1, \ldots, T\}$ such that $c_i^t(a) > 0$. Let t be the last period where this is the case, i.e., $c_i^\tau(a) = 0$ for all $\tau \in \{t+1, \ldots, T\}$ and for all $i \in I$. Now, consider the strategy b_i such that $c_i^{h'}(b_i, a_{-i}) = c_i^{h'}(a)$ at every $h' \notin H(b_i, a_{-i}) \cap \big(H_{t+1} \cup \cdots \cup H_T\big)$, and moreover $c_i^{h'}(b_i, a_{-i}) = 0$ at every $h' \in H(b_i, a_{-i}) \cap \big(H_{t+1} \cup \cdots \cup H_T\big)$, i.e., b_i contributes 0 at all (realized) histories that weakly follow h and agrees with a_i at all other histories. Then, observe that b_i is a profitable deviation from a_i given a_{-i}, since

$$
\begin{aligned}
u_i(a) &= 20 - \sum_{\tau=1}^{T} c_i^\tau(a) + \frac{r}{4}\sum_{\tau=1}^{T}\sum_{j=1}^{4} c_j^\tau(a) \\
&= 20 - \sum_{\tau=1}^{t} c_i^\tau(a) + \frac{r}{4}\sum_{\tau=1}^{t}\sum_{j=1}^{4} c_j^\tau(a) \\
&< 20 - \sum_{\tau=1}^{t} c_i^\tau(b_i, a_{-i}) + \frac{r}{4}\sum_{\tau=1}^{t}\sum_{j=1}^{4} c_j^\tau(b_i, a_{-i}) \\
&\leq 20 - \sum_{\tau=1}^{t} c_i^\tau(b_i, a_{-i}) + \frac{r}{4}\sum_{\tau=1}^{t}\sum_{j=1}^{4} c_j^\tau(b_i, a_{-i}) + \frac{r}{4}\sum_{\tau=t+1}^{T}\sum_{j=1}^{4} c_j^\tau(b_i, a_{-i}) \\
&= 20 - \sum_{\tau=1}^{T} c_i^\tau(b_i, a_{-i}) + \frac{r}{4}\sum_{\tau=1}^{T}\sum_{j=1}^{4} c_j^\tau(b_i, a_{-i}) \\
&= u_i(b_i, a_{-i}).
\end{aligned}
$$

Hence, $a_i \notin BR_i(a_{-i})$, thus contradiction our initial hypothesis that a is a NE. Therefore we conclude that there is no $t \in \{1, \ldots, T\}$ such that $c_i^t(a) > 0$, which completes the proof. ■

The following result is rather straightforward to prove, by applying – similarly to the part (i) of the previous proposition – the backward induction procedure.

Proposition 4 *Consider the public good game (with growth as defined above) with punishment. The unique SPE is such that every player contributes 0 at every contribution history and punishes 0 at every punishment history (both on and off the equilibrium path), i.e., if $a \in A$ is a SPE, then $c_i^h(a) = 0$ for every $i \in I$ and for all $h \in H^c$ and $p_i^h(a) = 0$ for every $i \in I$ and for all $h \in H^p$.*

Proof. Let $a \in A$ be an SPE. The proof proceeds by induction on t. In particular, we first prove our claim for T. Subsequently, we take an arbitrary $t \in \{1, \dots, T-1\}$, and we assume that our claim is true for every $\tau \in \{t+1, \dots, T\}$. Then, it suffices to prove our claim for t.

Initial step. Take an arbitrary $h \in H_T^p$ – not necessarily on the path induced by a – and assume that $p_{i,j}^h(a) > 0$ for an arbitrary pair $(i,j) \in I \times I$. Then, it follows directly that

$$\begin{aligned} u_i(a_i, a_{-i}|h) &= \min_{J \in \mathcal{J}_i^T} \left\{ W_i^T - \sum_{j \in J} 3p_{j,i}^h(a) \right\} - \sum_{j \neq i} p_{i,j}^h(a) \\ &< \min_{J \in \mathcal{J}_i^T} \left\{ W_i^T - \sum_{j \in J} 3p_{j,i}^h(b_i, a_{-i}) \right\} - \sum_{j \neq i} p_{i,j}^h(b_i, a_{-i}) \\ &= u_i(b_i, a_{-i}|h), \end{aligned}$$

with b_i being i's strategy that agrees with a_i at all $h' \neq h$, while $p_{i,j}^h(b_i, a_{-i}) = 0$. Indeed, notice that $p_{j,i}^h(a) = p_{j,i}^h(b_i, a_{-i})$ and $p_{i,j}^h(a) > 0 = p_{i,j}^h(b_i, a_{-i})$. This is because i's own punishment to j will be executed irrespective of the value of $\min_{J \in \mathcal{J}_i^T} \left\{ W_i^T - \sum_{j \in J} 3p_{j,i}^h(a) \right\}$.

Now take an arbitrary $h \in H_T^c$ – not necessarily on the path induced by a – and assume that $c_i^h(a) > 0$ for an arbitrary $i \in I$. The proof is almost identical to the one of the previous proposition. Indeed, take another strategy $b_i \in A_i$ agreeing with a_i at every $h' \neq h$, while $c_i^h(b_i, a_{-i}) = 0$. Given that the strategy profile $a \in A$ prescribes that no player punishes at any history in H_T^p, we obtain

$$\begin{aligned} u_i(a_i, a_{-i}|h) &= N_i^h - c_i^h(a_i, a_{-i}) + \frac{r}{4} \sum_{j=1}^{4} c_j^h(a_i, a_{-i}) \\ &< N_i^h + \frac{r}{4} \sum_{j=1}^{4} c_j^h(a_i, a_{-i}) \\ &= u_i(b_i, a_{-i}|h). \end{aligned}$$

Inductive step. Now fix an arbitrary $t \in \{1, \dots, T-1\}$, and we assume that for every $\tau \in \{t+1, \dots, T\}$, it is the case that (i) $c_i^h(a) = 0$ for all $h \in H_\tau^c$ and all $i \in I$, and (ii) $p_{i,j}^h(a) = 0$ for all $h \in H_\tau^p$ and all $(i,j) \in I \times I$.

Take an arbitrary $h \in H_t^p$ – not necessarily on the path induced by a – and assume that $p_{i,j}^h(a) > 0$ for an arbitrary pair $(i,j) \in I \times I$. Then, following exactly the same reasoning as in the initial step (and given the fact that according to the strategy profile a, every player will contribute 0 and will punish 0 at all histories that follow h), it will be the case that

$$u_i(a_i, a_{-i}|h) < u_i(b_i, a_{-i}|h),$$

with b_i being the strategy that agrees with a_i at all $h' \neq h$ while $p_{i,j}^h(b_i, a_{-i}) = 0$.

Finally, consider an arbitrary $h \in H_t^c$ – not necessarily on the path induced by a – and again assume that $c_i^h(a) > 0$ for an arbitrary $i \in I$. Then similarly to the initial step, take another strategy $b_i \in A_i$ agreeing with a_i at every $h' \neq h$, while $c_i^h(b_i, a_{-i}) = 0$. Given that the strategy profile $a \in A$ prescribes that no player contributes or punishes a positive amount at any history following h, we obtain

$$u_i(a_i, a_{-i}|h) < u_i(b_i, a_{-i}|h),$$

which completes the proof. ■

C.3 Reputation effects to predicted behavior

In their seminal paper, Kreps et al. (1982) showed that in a finitely repeated prisoner's dilemma, once a small grain of imperfect information is introduced, cooperation for a minimum number of periods is sustained as part of every sequential equilibrium (see also Kreps et al., 1982). In particular, in their setting they consider a player who – at the beginning of the game – assigns some small probability $\mu > 0$ to the event that the opponent will follow the tit-for-tat (henceforth, TFT) strategy, and maintains this belief unless it is contradicted by the actual path of play. Here we extend this idea to public good games (with and without growth). Let us first formally introduce the setting.

Tit-for-tat strategy. We define tit-for-tat (TFT) in the game without punishment as the strategy that begins with a (full) contribution of $a_i^{h_1} = N_i^{h_1}$ tokens at the initial history h_1, and then at every subsequent history h_t the player's proportional contribution (wrt to the endowment $N_i^{h_t}$ at that period) is as close as possible the minimum proportional contribution chosen by her opponents at the immediate predecessor h_{t-1}, i.e.,

$$a_i^{h_t} \in \arg \min_{a_i \in A_i^{h_t}} \left| \frac{a_i}{N_i^{h_t}} - \min_{j \neq i} \frac{a_j^{h_{t-1}}}{N_j^{h_{t-1}}} \right| .$$

Obviously, in the standard case without growth the previous definition yields $a_i^{h_t} = \min_{j \neq i} a_j^{h_{t-1}}$.

Information structure. We assume that at the beginning of the game, each player $i \in I$ believes with probability $\mu > 0$ that every opponent $j \neq i$ follows the TFT strategy and with probability $1 - \mu$ that every opponent is rational.[2] Then, at each history *that is consistent with all of her opponents having played according to TFT so far*, player i continues having the same beliefs.[3] On the other hand, if *at least one opponent has already deviated from TFT*, player i updates her beliefs, now assigning probability 1 to every $j \neq i$ being rational. Finally we assume that these beliefs are commonly believed.[4]

This informational context can be formally modelled as an incomplete information game, using a type-based model, as further developed by Battigalli and Siniscalchi (1999, 2002). Formally, for each player $i \in I$, there are two types $T_i = \{t_i^R, t_i^{TFT}\}$, viz., the rational type t_i^R whose payoff function at each history is the one given the standard public good game (with or without growth), and the TFT type t_i^{TFT} whose payoff function is such that TFT is a strictly dominant strategy. At every history h where all opponents of i have played in accordance to TFT at every preceding history, every $t_i \in T_i$ has beliefs described by the probability measure $\lambda_i^h(t_i) \in \Delta(T_{-i})$ which keeps assigning probability μ to $(t_j^{TFT})_{j \neq i}$ and probability $1 - \mu$ to $(t_j^R)_{j \neq i}$. On the other hand, at every history h where at least one opponent $j \neq i$ has deviated from TFT at some preceding history, every $t_i \in T_i$ has beliefs described by the probability measure $\lambda_i^h(t_i) \in \Delta(T_{-i})$ which attaches probability 1 to $(t_j^R)_{j \neq i}$. Note that in this framework, it is commonly believed at some history h that every player is rational, if for all $i \in I$ and for every $t_i \in T_i$ it is the case that $\lambda_i^h(t_i)\big((t_j^R)_{j \neq i}\big) = 1$. This is for instance the case at histories h where at least two players have deviated from the TFT strategy at preceding histories (see Observation 2 below).

Let us first make two rather straightforward preliminary observations.

Observation 1 *Fix an arbitrary history $h \in H$ where it is commonly believed that every player is rational. Then, it is commonly believed that every player contributes 0 from that history onwards.*

[2]We could have instead allowed i to form beliefs about each opponent independently. However, this would only make our analysis more complex without changing the qualitative nature of our results.

[3]The underlying idea is very similar to the one of strong belief, which is widely used in the characterization of forward induction in dynamic games (Battigalli and Siniscalchi, 2002). In particular, strong belief says that an event is believed as long as it is consistent by past observation.

[4]Kreps et al. (1982) use the term commonly known. This is due to the fact that at the early years of game theory "knowledge" was used for "probability 1 belief". Nowadays, it is standard to use the term "belief" and "common belief" instead.

The proof of this claim is identical to the one in Kreps et al. (1982, Step 1). In particular, if it is commonly believed that everybody's type is t_i^R, then it is commonly believed that a standard public good game is played, and by backward induction it follows that every player will choose 0 at every subsequent history, both on and off the equilibrium path.

Observation 2 *Consider a history $h \in H$ such that at least two players have deviated from the TFT strategy. Then, it is commonly believed that every player will contribute 0 from that history onwards.*

To see that this is the case, recall that when a player i deviates from the TFT strategy, then every $j \neq i$ believes (at every subsequent history) that every $k \neq j$ is rational, i.e., at h every $t_j \in T_j$ assigns probability 1 to $(t_k^R)_{k \neq j}$. Thus, if two players deviate from the TFT strategy then every player believes that everybody else is rational and this is commonly believed. Hence, from the previous observation, it becomes commonly believed that everybody will contribute 0 from that history onwards.

C.3.1 Standard public good game without growth

Now, suppose that we are in the standard setting without growth. Then, the following result shows that upon being observed that a player has chosen an action that is not consistent with the TFT strategy, it becomes commonly believed that everybody will contribute 0 from that point onwards.

Lemma 5 *Fix an arbitrary history $h \in H$ such that only player i has deviated from the TFT strategy. Then, it becomes commonly believed that every $j \in I$ will contribute 0 from that history onwards.*

Proof. If i has deviated from the TFT strategy at some history preceding h, every $j \neq i$ believes that every $k \neq j$ is rational, whereas i keeps assigning at h probability μ to the event that every $j \neq i$ is of type t_j^{TFT}. Obviously, every rational player will contribute 0 at every history in H_T – where T is the total number of rounds – and this is commonly believed. Now, consider some history $h_{T-1} \in H_{T-1}$ that follows h. First, notice that every $j \neq i$ believes that every $k \neq j$ is rational, and therefore will contribute 0 at every subsequent period. Hence, t_j^R will also contribute 0 at h_{T-1}, as she (correctly) believes her current action does not affect the opponents' future action. Now, let us turn to player i, and assume that no player other than i has deviated from TFT up to h_{T-1}, thus implying that i keeps attaching probability μ to the opponents' type profile being $(t_j^{TFT})_{j \neq i}$. Furthermore, let us assume that the t_j^{TFT} would contribute x at h_{T-1}. Then, i's expected payoff from choosing $y \leq x$ at h_{T-1} is equal to

$$U_i^{h_{T-1}}(y) = \mu\Big(\underbrace{20 - y + \frac{1.5}{4}y + \frac{1.5}{4}3x}_{\text{payoff at } T-1} + \underbrace{20 + \frac{1.5}{4}3y}_{\text{payoff at } T}\Big) + (1-\mu)\Big(\underbrace{20 - y + \frac{1.5}{4}y}_{\text{payoff at } T-1} + \underbrace{20}_{\text{payoff at } T}\Big),$$

which is maximized when $y = 0$, thus implying that the rational t_i^R will contribute 0. This also means that the TFT type t_j^{TFT} would also contribute 0 at every history in in H^T that follows h_{T-1}, as he will imitate i. Continue inductively to prove that at h it is commonly believed that every player will contribute 0. ■

Proposition 6 *Fix an arbitrary symmetric sequential equilibrium and let $(h_1, \ldots, h_T, z)$ be the equilibrium path. Then, there is some $t \in \{1, \ldots, T\}$, such that every rational player t_i^R contributes the full endowment N_i^h at the first t histories (i.e., at all $h \in \{h_1, \ldots, h_t\}$), and 0 at the remaining histories (i.e., at every $h \in \{h_{t+1}, \ldots, h_T\}$).*

Proof. Take a strategy profile such that every player contributes 20 at all histories up to history h_t and 0 all other histories following h_t as well as at all histories off this path. First, we show that this is a sequential equilibrium. For starters observe that off the path $(h_1, \ldots, h_T, z)$ every player is rational, and this is commonly believed, implying that each player's beliefs satisfy the requirements of a sequential equilibrium. Now, let us take an arbitrary history $h \in \{h_1, \ldots, h_T\}$. Notice that at h_t, each player i continues believing with probability μ that all $j \neq i$ are of type t_j^{TFT}. This is because, up to that history, no player has deviated from the TFT strategy. However, at every history following h_t the *rational player* will contribute 0, viz., both at h_{t+1} as well as off the path. Moreover, let $K_t := T - t + 1$ denote the number of periods remaining at a history in H_t, and therefore at our history h_t. Hence, by choosing any strategy that assigns a contribution $x < 20$ at h, player i's expected payoff becomes

$$U_i^{h_t}(x) \quad = \quad \underbrace{20 - x + \frac{1.5}{4}3 \cdot 20 + \frac{1.5}{4}x +}_{\text{payoff at } t} \qquad \underbrace{20(K_t - 1)}_{\text{payoff at remaining periods}}$$

Obviously, among all the possible deviations from TFT, the optimal one is to choose $x = 0$, in which case

$$U_i^h(0) = \frac{45}{2} + 20K_t.$$

On the other hand the TFT strategy induces an expected payoff of

$$\begin{aligned} U_i^{h_t}(TFT) &= \mu(1.5 \cdot 20K_t) + (1-\mu)\Big(\underbrace{1.5 \cdot 20}_{\text{payoff at } t} + \underbrace{\frac{1.5}{4}20}_{\text{payoff at } t+1} + \underbrace{20(K_t - 2)}_{\text{payoff at remaining periods}}\Big) \\ &= \frac{5}{2}(\mu - 1) + (10\mu + 20)K_t. \end{aligned}$$

Hence, $U_i^h(TFT) > U_i^h(0)$ if and only if

$$K_t > \hat{K}_t(\mu) := \frac{50 + 5\mu}{20\mu}, \tag{8}$$

where $\hat{K}_t(\mu)$ is the least number of remaining periods so that players continue to contribute. ■

C.3.2 Public good game with growth

Now, we are going to show that the main conclusion of Proposition 6 continues holding in public good games with growth, i.e., the structure of sequential equilibria is the same. However, what changes is the lower bound on the number of periods that the players contribute their whole endowment, i.e., *cooperation lasts longer*. For computation simplicity, we are going to focus on cases where $\mu < 10/35$. This is a rather mild assumption, as this entire literature restricts attention to very small μ's (e.g., see Kreps et al., 1982).

Lemma 7 *Let $\delta < 10/35$ and fix an arbitrary history $h \in H$ such that only player i has deviated from the TFT strategy. Then, it becomes commonly believed that every rational player will contribute 0 from that history onwards.*

Proof. The proof follows similar steps as the one of Lemma 5 above. The difference is that at some history in $h \in H_{T-1}$ where only player i has deviated up to that point, i's expected payoff as a function of i's proportional contribution $\beta \in [0, 1]$ is

$$\begin{aligned} U_i^h(\beta) &= \mu\Big(\underbrace{N_i^h}_{\text{endowment at } T-1} \underbrace{- \beta N_i^h + \frac{1.5}{4}\beta N_i^h + \frac{4.5}{4}\alpha N_j^h}_{\text{payoff at } T-1} + \underbrace{\frac{4.5}{4}\beta\big(N_j^h - \alpha N_j^h + \frac{1.5}{4}\beta N_i^h + \frac{4.5}{4}\alpha N_j^h\big)}_{\text{payoff at } T}\Big) \\ &+ (1-\mu)\Big(N_i^h - \beta N_i^h + \frac{1.5}{4}\beta N_i^h\Big) \\ &= \mu\Big(N_i^h - \frac{2.5}{4}\beta N_i^h + \frac{4.5}{4}\alpha N_j^h + \frac{4.5}{4}\beta\big(N_j^h + \frac{0.5}{4}\alpha N_j^h + \frac{1.5}{4}\beta N_i^h\big)\Big) + (1-\mu)\big(N_i^h - \frac{2.5}{4}\beta N_i^h\big) \end{aligned}$$

with $\alpha \in [0, 1]$ denoting the proportional contribution of every $j \neq i$ at h. Then, notice that, given our condition on δ, the expected payoff $U_i^h(\beta)$ is maximized for $\beta = 0$, irrespective of α. To see this, we differentiate $U_i^h(\beta)$ wrt to β, and then using the facts that $\alpha \leq 1$ and $\beta \leq 1$ and $N_j^h < N_i^h$, we obtain

$$\frac{\partial U_i^h}{\partial \beta} < \mu\Big(\frac{2}{4}N_j^h + \frac{2.25}{16}N_j^h + \frac{13.5}{16}N_j^h\Big) - (1-\mu)\frac{2.5}{4}N_j^h$$

which is in turn negative if $\mu < 10/35$. Hence, i will contribute 0 at $h \in H_{T-1}$. This implies that at the last history both t_j^R as well as t_j^{TFT} will contribute 0. Then, by working backwards we inductively prove that every player will contribute 0 at all histories following the first deviation of i, and this is commonly believed. ■

Proposition 8 *Let $\delta < 10/35$. Fix an arbitrary symmetric sequential equilibrium and let $(h_1, \dots, h_T)$ be the equilibrium path. Then, there is some $t \in \{1, \dots, T\}$, such that every rational player t_i^R contributes the full endowment N_i^h at the first t histories, i.e., at all $h \in \{h_1, \dots, h_t\}$, and 0 at the remaining histories, i.e., at every $h \in \{h_{t+1}, \dots, h_T\}$.*

Proof. The proof of this claim is almost identical to the one of Proposition 6 above. In particular, first notice that at h_t, each player i continues believing with probability μ that all $j \neq i$ are of type t_j^{TFT}. This is because, up to that history, no player has deviated from the TFT strategy. Hence, by choosing any strategy that deviates from contributing $\beta = 1$ at h_t, player i's expected payoff becomes

$$U_i^{h_t}(0) = N_i^{h_t} + \frac{1.5}{4} 3 N_i^{h_t}.$$

On the other hand the TFT strategy induces an expected payoff of

$$U_i^{h_t}(TFT) = \mu\big(1.5^{K_t} N_i^{h_t}\big) + (1-\mu)\big(\frac{1.5^2}{4} N_i^{h_t}\big).$$

Hence, we obtain $U_i^{h_t}(TFT) > U_i^{h_t}(0)$ whenever it is the case that

$$K_t > \hat{K}_t^G(\mu) := \log_{1.5} \frac{6.25 + 2.25\mu}{4\mu},$$

where $\hat{K}_t(\mu)$ is the least number of remaining periods so that players continue to contribute. ■

D Additional Tables and Figures

This section contains additional tables and figures. Table 7 summarizes the number of independent observations, participants and sessions in all our treatments. Table 8 shows the order of sessions for our main treatments.

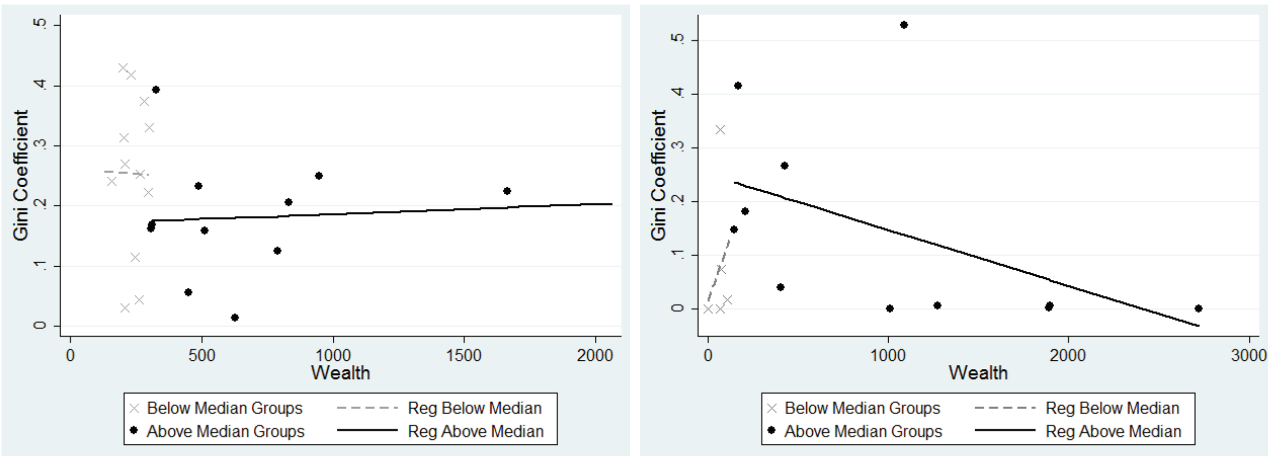
	15 periods	10 periods	Overall
W/o Punishment (NOPUNISH)	15 (60,2)	23 (92,3)	38 (152,5)
With Punishment (PUNISH)	15 (60,2)	21 (84,3)	36 (144,5)
No Inequality w/o punish (NOPUNISH-NOINEQUALITY)	-	24 (96,3)	24 (96,3)
No Inequality with punish (PUNISH-NOINEQUALITY)	-	14 (56,3)	14 (56,3)
No Growth w/o punish (NOPUNISH-NOGROWTH)	-	29 (116,4)	29 (116,4)
No Growth with punish (PUNISH-NOGROWTH)	-	23 (92,3)	23 (92,3)

Table 7: Number of Independent Observations (Participants, Sessions).

	Length	Punish	Groups
24/09/2012 11:30	15	NOPUNISH	7
24/09/2012 13:30	15	PUNISH	8
24/09/2012 15:30	15	NOPUNISH	8
05/10/2012 11:00	10	PUNISH	7
05/10/2012 13:30	10	NOPUNISH	8
05/10/2012 16:00	10	PUNISH	7
02/11/2012 11:00	15	PUNISH	7
02/11/2012 13:00	10	NOPUNISH	8
02/11/2012 14:30	10	NOPUNISH	7
02/11/2012 16:00	10	PUNISH	7

Table 8: Order of sessions for the main treatments. Sessions for the additional treatments reported on in Section 5.2 were conducted between 23/04/2014 and 25/06/2014.

Table 9 shows random effects OLS regressions of wealth on period and treatment dummy as well as interactions for the 10-period games. Table 10 shows the same analysis for the 15-period games. Tables 11 and 12 focus on inequality (Gini coefficients) as outcome. Figure D.1 shows the correlation between the wealth and Gini in period 10 only for the 23 (21) groups in NOPUNISH (PUNISH).



(a) NOPUNISH (b) PUNISH

Figure D.1: Correlation between the wealth and Gini in period 10 only for the 23 (21) groups in NOPUNISH (PUNISH). Each point represents one group with their period 10 wealth and Gini coefficient. Dots are above median groups and crosses below median groups. Lines are fitted values from linear regression of Gini on wealth. In the graph for treatment PUNISH some below median groups with a Gini coefficient of 1 are omitted from the graph (not the regression) for expositional clarity.

	Wealth					
	(1)	(2)	(3)	(4)	(5)	(6)
period (β_1)	40.50***	18.84***	45.06***			
	(7.59)	(7.23)	(9.27)			
PUNISH (β_2)	-80.98*	-76.47***		-28.44	-140.60***	61.57
	(48.64)	(25.83)		(57.24)	(10.67)	(79.21)
period × PUNISH (β_3)	9.55	-69.17***				
	(19.09)	(18.37)				
period^2 (β_4)		1.96				
		(1.27)				
period^2 × PUNISH (β_5)		7.15**				
		(3.19)				
Constant (α)	14.19	57.51***	-24.47	236.90***	170.90***	308.90***
	(20.26)	(9.13)	(24.39)	(21.85)	(6.68)	(33.72)
Test $\beta_1 + \beta_3 = 0$	50.05***	-50.33***				
p-value	0.0043	0.0029				
Test $\beta_4 + \beta_5 = 0$		9.111***				
p-value		0.0018				
Observations	440	440	440	440	220	220
Groups	44	44	44	44	22	22
Sample	All	All	All	All	below median	above median

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Random effects OLS regression of wealth on period and treatment dummy. Significance at the 1,5,10 percent level is denoted by ***,** ,*, respectively. Standard errors account for autocorrelation and are clustered at the matching group level. 10 period games only.

	Wealth					
	(1)	(2)	(3)	(4)	(5)	(6)
period (β_1)	95.66**	-57.83	67.59***			
	(47.62)	(56.89)	(25.92)			
PUNISH (β_2)	149.80	-49.10		-299.20	-128.20***	-401.70
	(212.50)	(112.40)		(195.40)	(24.01)	(325.30)
period × PUNISH (β_3)	-56.12	14.07				
	(50.80)	(64.19)				
period^2 (β_4)		9.59				
		(6.51)				
period^2 × PUNISH (β_5)		-4.38				
		(7.12)				
Constant (α)	-242.7	192.2**	-167.8	522.6***	175.7***	826.1***
	(198.80)	(97.08)	(107.10)	(182.50)	(12.98)	(307.7)
Test $\beta_1 + \beta_3 = 0$	39.54**	-43.76				
p-value	0.0253	0.1409				
Test $\beta_4 + \beta_5 = 0$		5.21*				
p-value		0.0709				
Observations	450	450	450	450	225	225
Groups	30	30	30	30	15	15
Sample	All	All	All	All	below median	above median

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Random effects OLS regression of wealth on period and treatment dummy. Significance at the 1,5,10 percent level is denoted by ***,** ,*, respectively. Standard errors clustered at the matching group level. 15 period games only.

	Gini coefficient					
	(1)	(2)	(3)	(4)	(5)	(6)
period (β_1)	0.017***	0.053***	0.012***			
	(0.002)	(0.009)	(0.003)			
PUNISH (β_2)	0.050**	0.084***		-0.010	0.038	-0.049**
	(0.020)	(0.030)		(0.036)	(0.061)	(0.024)
period × PUNISH (β_3)	-0.011*	-0.028				
	(0.006)	(0.018)				
period^2 (β_4)		-0.003***				
		(0.000)				
period^2 × PUNISH (β_5)		0.001				
		(0.001)				
Constant (α)	0.063***	-0.008	0.087***	0.161***	0.191***	0.129***
	(0.007)	(0.012)	(0.010)	(0.017)	(0.026)	(0.018)
Test $\beta_1 + \beta_3 = 0$	0.006	0.025*				
p-value	0.2609	0.0993				
Test $\beta_4 + \beta_5 = 0$		-0.002				
p-value		0.3002				
Observations	440	440	440	440	220	220
Groups	44	44	44	44	22	22
Sample	All	All	All	All	below median	above median

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11: Random effects OLS regression of Gini coefficient on period and treatment dummy. Significance at the 1,5,10 percent level is denoted by ***,** ,*, respectively. Standard errors account for autocorrelation and are clustered at the matching group level. 10 period games only.

	Gini coefficient					
	(1)	(2)	(3)	(4)	(5)	(6)
period (β_1)	0.005**	0.015*	0.002			
	(0.002)	(0.009)	(0.002)			
PUNISH (β_2)	0.033	0.083*		-0.019	-0.017	-0.025
	(0.039)	(0.044)		(0.029)	(0.051)	(0.030)
period × PUNISH (β_3)	-0.006	-0.024**				
	(0.005)	(0.011)				
period^2 (β_4)		-0.000				
		(0.000)				
period^2 × PUNISH (β_5)		0.001				
		(0.000)				
Constant (α)	0.087***	0.058***	0.104***	0.134***	0.145***	0.124***
	(0.011)	(0.017)	(0.020)	(0.021)	(0.038)	(0.023)
Test $\beta_1 + \beta_3 = 0$	-0.001	0.015				
p-value	0.8656	0.2464				
Test $\beta_4 + \beta_5 = 0$		0.001				
p-value		0.3208				
Observations	450	450	450	450	225	225
Groups	30	30	30	30	15	15
Sample	All	All	All	All	below median	above median

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Random effects OLS regression of Gini coefficient on period and treatment dummy. Significance at the 1,5,10 percent level is denoted by ***,** ,*, respectively. Standard errors clustered at the matching group level. 15 period games only.

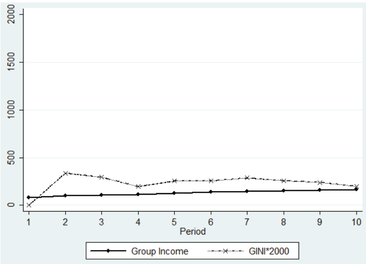
	(1) norm contribution	(2) contribution
PUNISH	11.27	8.057***
	(11.89)	(1.135)
NOPUNISH-NOGROWTH	-23.77***	
	(8.21)	
PUNISH-NOGROWTH	-23.52***	
	(8.21)	
NOPUNISH-NOINEQUALITY		26.29***
		(8.632)
PUNISH-NOINEQUALITY		50.74***
		(13.83)
Constant	24.09***	31.46***
	(8.21)	(10.99)
Observations	5,080	5,376
Number of Participants	448	504

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

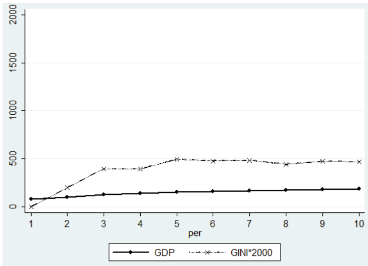
Table 13: (Normalized) contributions regressed on treatment dummies. Simple OLS regression. Standard errors clustered by matching group. Baseline is treatment NOPUNISH. ***,**,* significance at 1,5,10 percent level.

E Matching Group Figures

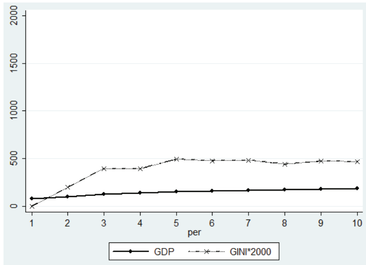
Figures E.1-E.2 show the evolution of wealth and Gini coefficient over time for the six poorest and six richest matching groups in each treatment as measured by period 10 wealth. Graphs on additional matching groups are available upon request. In NOPUNISH (Figure E.1) the evolution of both indicators is relatively smooth. In PUNISH (Figure E.2) an interesting phenomenon can be observed. In groups where the Gini coefficient rises sharply in early periods (e.g. groups 201 or 208), there is so much punishment that wealth ends up being zero.



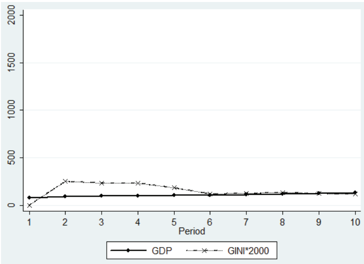
(a) Group 101



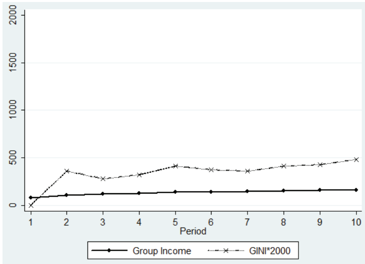
(b) Group 302



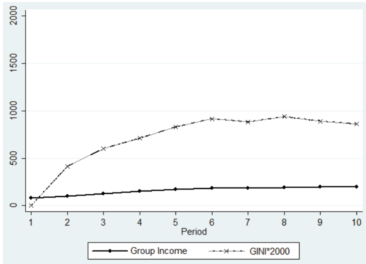
(c) Group 303



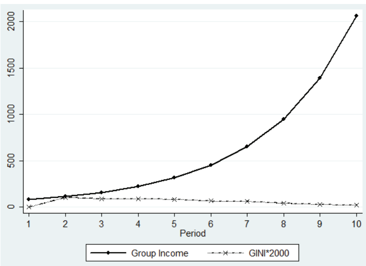
(d) Group 308



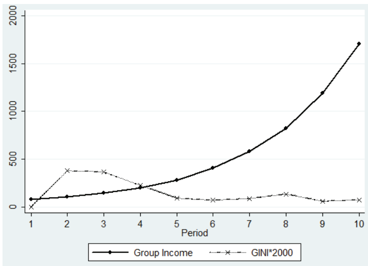
(e) Group 804



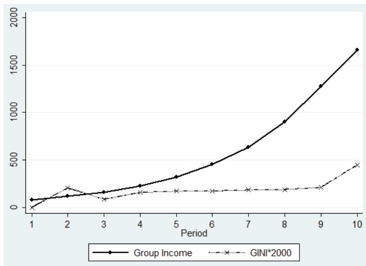
(f) Group 907



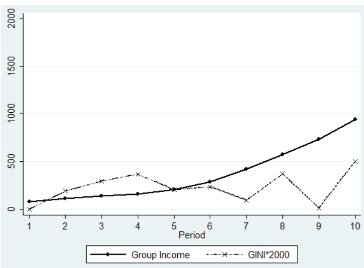
(g) Group 107



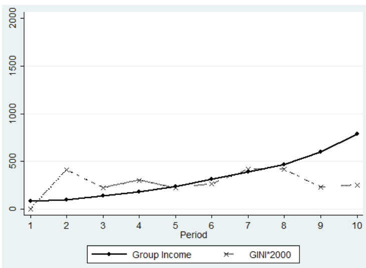
(h) Group 304



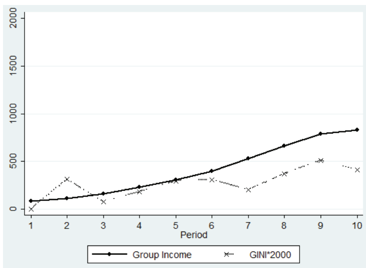
(i) Group 502



(j) Group 504

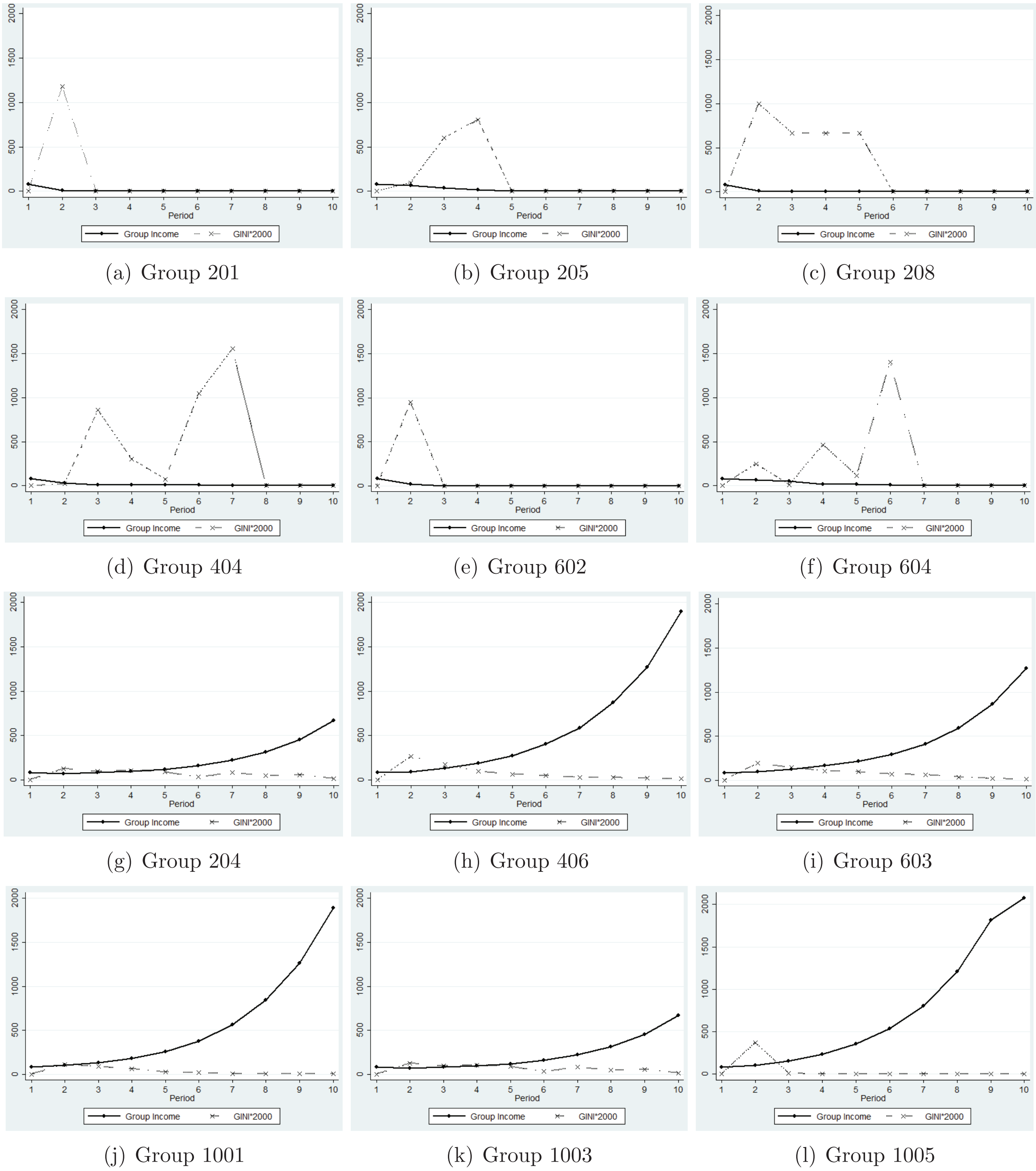


(k) Group 901



(l) Group 904

Figure E.1: Wealth and Gini coefficient across the six poorest (panels (a)-(f)) and six richest ((g)-(l)) matching groups (as measured by $t = 10$ wealth). Treatment NOPUNISH. Gini coefficient is multiplied by 2000 to be on the same scale as wealth.



(a) Group 201 (b) Group 205 (c) Group 208

(d) Group 404 (e) Group 602 (f) Group 604

(g) Group 204 (h) Group 406 (i) Group 603

(j) Group 1001 (k) Group 1003 (l) Group 1005

Figure E.2: Wealth and Gini coefficient across the six poorest (panels (a)-(f)) and six richest ((g)-(l)) matching groups (as measured by $t = 10$ wealth). Treatment PUNISH. Gini coefficient is multiplied by 2000 to be on the same scale as wealth.

F Questionnaire Data

In this section we summarize the results from our post-experimental questionnaire. All questions can be found in Online Appendix B. Before we discuss the responses we should mention that due to computer problems in some of the sessions our questionnaire data are incomplete. Those problems were exogenous to session and participant characteristics, so our collected data should be representative. However, the reader should be aware that they don't contain all our participants. We were able to collect full questionnaires from 124 out of 152 participants in treatment NOPUNISH and 84 out of 144 participants in treatment PUNISH. Table 14 summarizes the key characteristics of our participants. About half of them are female. Also about half are German, but there are also significant percentages of Dutch, Western European (Belgium, Luxemburg, France or UK) and Eastern European participants. They are on average 21.5 years old in both treatments. The youngest participant was 17 and the oldest 35. Around 40 percent of them are business students and almost all others students from other fields (very few non-students). They have spent on average 2 years at university. Our risk aversion measure has full support in our sample and there are no significant treatment differences in the distributions of any of the variables mentioned in Table 14.

	NOPUNISH	PUNISH
Gender (Share female)	0.42	0.52
Share German	0.56	0.43
Share Dutch	0.13	0.10
Share BEL/LUX/FRA/UK	0.11	0.13
Share Eastern Europe	0.10	0.20
Average Age (Range)	21.5(18, 35)	21.5(17, 28)
Share Business	0.41	0.40
Share Economics	0.20	0.12
Share European Studies	0.07	0.12
Share Psychology	0.08	0.05
Years studied (Range)	2.1(0, 10)	2.0(0, 5)
Risk Aversion (Range)	3.39(0, 7)	3.19(0, 7)

Table 14: Summary Statistics Questionnaire Data. Only Nationality Categories and Fields of Study with more than 10 percent answers are mentioned explicitly. The variable risk aversion can take values from 0 to 7, where 0 is most risk averse and 7 least risk averse.

Table 15 summarizes the responses to the personality questionnaire. Again the distribution of answers is very similar across treatments.

	NOPUNISH	PUNISH
Q1 I am a quick thinker	5.18	5.40
Q2 I get easily offended	3.58	3.66
Q3 very satisfied	5.07	5.16
Q4 very dependent	2.67	2.75
Q5 generally happy	5.71	5.75
Q6 work important	4.77	4.89
Q7 family important	5.67	6.03
Q8 friends important	6.01	6.05
Q9 religion important	2.47	2.26
Q10 politics important	3.65	3.60
Q11 most people trusted	3.72	3.88
Q12 hard work better	5.48	5.44
Q13 government responsible	4.29	4.45
Q14 incomes equal	3.78	3.76

Table 15: Summary Statistics Questionnaire Data, Mean Reply to Personality Characteristics Questions of the form "How strongly do you agree to the following statements?" 1 - disagree strongly, 7 - agree strongly. The exact statements can be found in Online Appendix B.

We then regress our measures of growth (wealth) and inequality (Gini) in the two treatments on the questionnaire data. We use simple OLS regressions of wealth and Gini in period 10 on individual questionnaire data and we cluster standard errors by matching group. Table 17 shows the results for treatment NOPUNISH. Overall our questionnaire measures have a hard time to explain the variation in wealth and Gini and almost all of them are insignificant. There might be somewhat of a gender effect in treatment NOPUNISH. In particular wealth seems to be lower in groups with more women. Strong agreement to the statement “Friends play an important role in my life” seems to predict somewhat higher wealth in treatment NOPUNISH. Both of these results should be interpreted with care, though, since we regress on quite a large set of variables. The overall message seems to be that our questionnaire data *cannot* explain the variation in wealth and Gini coefficient.

	NOPUNISH	PUNISH
above median wealth	2.31	1.70
	(2.18)	(1.42)
below median wealth	2.24	1.46
	(1.53)	(3.17)

Table 16: Average Donation (Std. Dev.) in Euros to Medics without Borders.

Table 18 shows the results of the analogous regression for treatment PUNISH. Here the result is even clearer. None of the variables seems systematically able to explain any of the variation in wealth or Gini observed in this treatment. There is a significant coefficient on risk aversion, indicating that higher risk aversion of group members might lead to higher wealth in these treatments. This effect would be intuitive if risk averse participants react more strongly to the threat of punishment, but it disappears once we stop controlling for the personality characteristics.

Finally we have a look at how much our participants decide to donate to Medics without Borders. Table 16 shows the average donation in Euros to medics without Borders. Participants in treatment PUNISH seem to donate somewhat less than participants in treatment NOPUNISH. We compare the distribution of donations using a two-sided ranksum test where we treat each individual donation as an independent observation. The two treatments are significantly different ($p = 0.0432$) on aggregate and if we restrict to below median groups ($p = 0.0134$), but not restricted to above median groups ($p = 0.4711$).

More interestingly, though, participants from groups with wealth above the median do *not* seem to contribute more on average than those from groups with below median wealth. There is no significant difference in treatment NOPUNISH ($p = 0.9195$) and a marginally significant difference in treatment PUNISH ($p = 0.0506$).

This is despite the fact that participants from groups with above median wealth earn 178 tokens on average in period 10 (189 in treatment PUNISH), while those from groups with below median wealth earn only 56 tokens (23 tokens) on average in period 10. This evidence suggests hence that participants in groups with above median wealth are *not* per se more altruistic than others.

	(wealth)	(wealth)	(Gini)	(Gini)
gender	−147.55**	−151.23**	−0.04*	−0.02
	(70.58)	(58.84)	(0.02)	(0.02)
age	−21.46	−17.36	−0.00	−0.00
	(21.89)	(19.57)	(0.00)	(0.00)
risk aversion	−38.45	−35.06	0.00	0.00
	(26.31)	(23.62)	(0.01)	(0.01)
Q1	−29.48		−0.00	
	(19.14)		(0.01)	
Q2	−2.33		0.00	
	(15.53)		(0.00)	
Q3	−8.93		−0.00	
	(42.65)		(0.01)	
Q4	−10.22		0.00	
	(24.77)		(0.00)	
Q5	28.45		0.02	
	(41.43)		(0.01)	
Q6	−6.60		0.00	
	(26.88)		(0.00)	
Q7	−30.41		−0.00	
	(25.03)		(0.00)	
Q8	96.04**		−0.00	
	(38.20)		(0.01)	
Q9	−15.25		0.00	
	(25.27)		(0.00)	
Q10	46.12		−0.02*	
	(28.45)		(0.01)	
Q11	11.35		0.00	
	(16.87)		(0.00)	
Q12	3.88		−0.00	
	(24.16)		(0.00)	
Q13	−60.60		0.00	
	(42.82)		(0.01)	
Q14	8.16		−0.00	
	(12.98)		(0.00)	
constant	43100.80	35194.95	3.91	1.02
	(43656.74)	(39033.02)	(10.07)	(8.02)
Observations	124	124	124	124
Groups	31	31	31	31
R^2	0.1387	0.0607	0.0780	0.0110
VCE robust S.E.	Yes	Yes	Yes	Yes

Table 17: OLS regression of period 10 wealth and Gini coefficient on questionnaire characteristics. Treatment NOPUNISH

	(wealth)	(wealth)	(Gini)	(Gini)
gender	−40.17	−47.60	−0.04	−0.03
	(169.29)	(204.07)	(0.04)	(0.05)
age	−7.17	0.10	0.01	0.01
	(26.43)	(25.60)	(0.01)	(0.02)
risk aversion	124.04**	73.43	0.00	0.00
	(58.87)	(47.51)	(0.02)	(0.02)
Q1	−2.91		0.01	
	(62.92)		(0.01)	
Q2	−40.90		0.02	
	(28.99)		(0.02)	
Q3	−30.15		−0.00	
	(50.43)		(0.01)	
Q4	−52.27		−0.02	
	(102.48)		(0.02)	
Q5	−67.12		−0.01	
	(68.65)		(0.02)	
Q6	117.22		−0.00	
	(75.48)		(0.01)	
Q7	5.52		0.02	
	(81.70)		(0.02)	
Q8	−53.67		−0.00	
	(61.58)		(0.01)	
Q9	74.55		0.01	
	(55.88)		(0.01)	
Q10	−14.53		−0.01	
	(40.04)		(0.01)	
Q11	34.14		0.00	
	(41.89)		(0.02)	
Q12	−53.44		0.00	
	(107.01)		(0.02)	
Q13	−41.32		0.01	
	(34.60)		(0.01)	
Q14	3.71		−0.02	
	(81.51)		(0.02)	
constant	15173.64	67.15	−20.02	−28.47
	(52764.84)	(50911.71)	(35.82)	(40.83)
Observations	84	84	84	84
Groups	21	21	21	21
R^2	0.1430	0.0243	0.1213	0.0352
VCE robust S.E.	Yes	Yes	Yes	Yes

Table 18: OLS regression of period 10 wealth and Gini coefficient on questionnaire characteristics. Treatment PUNISH