Essays in Applied Microeconomics

Inauguraldissertation
zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von

Mikhail Ananyev

aus Kolomna/UdSSR

2021

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferent:	Prof. Dr. Sebastian Kube
Zweitreferent:	Prof. Dr. Thomas Dohmen

Tag der mündlichen Prüfung: 18. Mai 2021

Acknowledgments

It is impossible for me to give enough thanks to the two of my supervisors, Prof. Sebastian Kube and Prof. Thomas Dohmen. They always encouraged me to take the next step whenever I fear moving forward, they guided me through the depths of economic theory, enlightened me with their understanding of the matters I studied, and supported me in every way possible. Their energy, humor, passion about research, and openness for discussion more than anything else inspired me and motivated me to move forward, search for answers and try new approaches.

I must thank my third supervisor, Prof. Hartmut Lehmann, for his reasonable critique and invaluable conversations, wich helped me to better understand what I want to do next.

I would like to thank my family and Nadezhda Zhuravleva, with whom we shared both joyful and poor times over the past years, for all their love and support.

I thank BGSE, the Institute of Labor Economics (IZA), the Institute of Applied Microeconomics (IAME), and Deutscher Akademischer Austauschdienst (DAAD) for their generous financial support.

I would like to thank Ms. Silke Kinzig at BGSE and Ms. Lydia Simons at IZA, who were able to seemingly magically find a solution to any non-academic problems I had been facing. Special thanks go to Ms. Andrea Reykers, Ms. Britta Altenburg at the BGSE, Ms. Sarah Assili at IZA, and Ms. Bénédicte Pillot-Bechtold at the Institute for International Economic Policy.

I would like to thank all the people from IZA, especially the publication team, with whom we had wonderful times in the office and after work. I am thankful to my friends at BGSE and IZA, including but not limited to Carl Heese, Axel Wogrolly, Branislav Albert, Marina Khismatullina, Lucas ter Steege, and Lars Meierwisch, who helped not only in the study and academic research, but also in daily life. Many thanks to Radost Holler and Raphael Suchy for their help with the lab.

Contents

List of Figures

List of Tables

Chapter I

Leading-by-Example in a Public Goods Game with Heterogeneous Returns*

I.1 Introduction

Individuals are frequently involved in collective activities that are beneficial for the group and its members (e.g. fund raising, scientific collaboration, or emissions reduction), giving rise to social dilemma problems. The literature has established that institutions play one of the central roles in mitigating social dilemma, mainly focusing on symmetric groups of identical players. However, in practice it is difficult to assume such similarities among group members. Individuals can differ not only in terms of the endowments that they possess, but also in their benefits from public goods (for example, they have different opportunity costs, individual cost savings from reduced CO_2 emissions, or additional private benefits from working on a joint project). These differences may induce trade-offs between various contribution rules that different players perceive as fair, undermining the effectiveness of established institutions. The question is to what extent the performance of these institutions depends on asymmetries among involved participants.

Leading-by-example has been shown to be an institution that serves as a coordination device fostering cooperation (see Frackenpohl et al., 2016; Gächter et al., 2012; Güth et al., 2007). It transforms the game into a sequential one: one person makes a contribution decision first (the leader), and the others observe this decision and make their decisions afterwards (the followers). The effect of this institution has been widely studied in symmetric groups, in which equal contributions lead to equal payoffs. However, little is known about how leading-by-example performs in asymmetric groups.[1] The main problem with asymmetric groups is that there can be less uniform perceptions of the equality of contributions, the equality of payoffs, and efficiency. Moreover, leaders with varying returns can lead differently, which in turn can be perceived differently by followers with

*I would like to thank Sebastian Kube, Simone Quercia, Thomas Dohmen and Felix Albrecht for their very helpful comments; financial support from the Deutscher Akademischer Austauschdienst (DAAD) and the Institute of Applied Microeconomics (IAME) is gratefully acknowledged.

[1] Au and Chung (2007); Levati et al. (2007); Neitzel and Sääksvuori (2013) study sequential move structure in groups with heterogeneous endowments. Ibanez and Schaffland (2018) study leadership in groups with heterogeneous returns, focusing on expectation about the performance of in-group versus out-group leaders which affects followers' returns from the public good. My focus is on the effect of varying the composition of returns in the group.

heterogeneous returns.

I investigate the effect of leading-by-example in a sequential linear public goods game with asymmetric groups of three people. The group members differ in their returns from the public good, whereby some benefit more (with high returns, H-players) than others (with low returns, L-players). The leader is a randomly-selected player who contributes first and whose contribution becomes publicly known to other members of the group. The return of the leader (high or low) differs between treatments. Followers always have different returns: one L-follower and one H-follower (allocated randomly). I study followers' reactions to the contribution decisions of leaders with different returns.

I innovate on existing experimental designs to overcome a number of problems, including an important selection issue with respect to the observed contributions of the leaders, namely that the distribution of the leader's contributions may differ depending on leaders' high or low returns.[2] This selection problem does not allow for the clear comparison of followers' reactions to the observed leaders' contributions. To solve this problem, I measure followers' cooperation preferences, namely conditional contributions with respect to all possible contributions of the leader (see Fischbacher and Gächter, 2010). This approach eliminates not only possible differences in the distributions of the leader's contributions but also the distortions in followers' reactions related to prior beliefs about the leader's contribution (see Selten, 1967).[3]

The results suggest that individual contribution preferences (conditional contributions) are systematically affected by the heterogeneity in returns from the public good among group members. I find that followers are more likely to follow a leader who benefits less from the public good than one who benefits relatively more, regardless of their own returns. These novel results provide further evidence that leaders serve as a coordination device, even in groups with asymmetric returns. Moreover, I provide evidence that such a coordination device can help to foster cooperation despite inequality considerations and the presence of multiple plausible contribution rules, whereby groups in which leaders have low returns better sustain cooperation (at a nearly efficient level) when it comes to a dynamic interaction in repeated games. These findings might be particularly interesting to managers and policy-makers who deal with belief management and group insurance.

The remainder of the paper is organized as follows. I discuss previous literature guiding this research in Section I.2 and present the design of the experiment in Section I.3. I derive behavioral predictions on the effect of leading-by-example in groups with heterogeneous benefits from the public good in Section I.4. The results are presented in Section I.5, before Section I.6 presents the results from additional treatments for repeated simultaneous and repeated sequential games. Finally, Section I.7 concludes.

[2] Technically speaking, the supports of the distributions might differ.

[3] The experimental approach also allows clearly identifying the treatment effect and controlling for the information structure. It also allows eliciting players' beliefs and measuring individual characteristics that are typically difficult to be observed in the field, such as trust and risk preferences.

I.2 Literature Review

In collective activities such as raising funds or scientific collaboration, individuals can benefit from the group performance regardless of their own contribution, namely they can have incentives to free-ride despite the fact that the most beneficial outcome for all members would be to cooperate. This social dilemma is captured in a public goods game, in which a Pareto efficient outcome — when all players contribute as much as possible — is a dominated one because it is always profitable to free-ride, i.e. to contribute zero to the public good. However, it has been widely documented that individuals do not behave as selfish profit-maximizing agents, as people contribute positive amounts to the public good, even in one-shot games.[4] One explanation for such behavior is that people actually care about fairness and other issues related to their *relative* payoffs, which is captured by social preferences.[5] The introduction of social preferences can change the game from a social dilemma into a coordination problem with many equilibria (see, for example, Hartig et al., 2015, for a detailed comparison of models of inequity aversion).

The existing literature on public goods games has suggested various mechanisms that deal with free-rider and coordination problems. Some mechanisms change the incentive schemes to make group members behave in the group's interests (Andreoni and Gee, 2012; Fehr and Gächter, 2000, 2002; Gürerk et al., 2006), while some change the environmental (Dufwenberg et al., 2011; Kube et al., 2015; Van Dijk and Wilke, 1995, 2000) or information structure (Erev and Rapoport, 1990; Fischbacher and Gächter, 2010; Gächter and Thöni, 2005; Hartig et al., 2015; Levy et al., 2011). However, it is debatable whether some mechanisms are more efficient in promoting cooperation than others. While changing the incentive schemes — for example, introducing punishment for non-cooperative behavior — seems to stimulate higher contributions, the punishment typically costs players part of their payoffs, making this mechanism costly. Leadership typically transforms the game into a sequential one, although without enforcement (e.g. ostracism) it may lack the power to mitigate social dilemmas. Furthermore, its effectiveness hinges upon players having social preferences and not being purely profit-maximizing agents.

This paper builds on previous studies on public goods games, bringing together two strands of the literature on asymmetries between players in public goods experiments as well as the literature on leadership.

I.2.1 Literature on Leadership

Leadership has been shown to be an institution that improves cooperation (see Frackenpohl et al., 2016; Gächter et al., 2012; Güth et al., 2007), as it serves as a coordination mechanism. It has been found to successfully promote cooperation, especially when leaders

[4] See Ledyard (1995) for an overview of early experimental evidence.

[5] There are different models of social preferences, like models for inequity aversion (see Bolton and Ockenfels, 2000; Cappelen et al., 2007; Charness and Rabin, 2002; Fehr and Schmidt, 1999) and models of reciprocity (e.g. Falk and Fischbacher, 2006).

are granted some authority to punish free-riding (e.g. Güth et al., 2007), reward cooperation (Gürerk et al., 2009), or choose to reallocate revenues (Van der Heijden et al., 2009). Even in the absence of the power of authority, leaders can have a positive effect on the contributions of other players, as they can help to induce social norms by displaying their own contribution to the other players (see Centorrino et al., 2013; Frackenpohl et al., 2016; Glöckner et al., 2011; Güth et al., 2007; Haigner and Wakolbinger, 2010),[6] signaling the quality of the public good (Potters et al., 2007), or directly suggesting a contribution norm (Levy et al., 2011). In the simplest case simply involving a sequential order of moves (called leading-by-example), one person makes a decision first (the leader), while the others (the followers) observe this decision and contribute afterwards.

Leading-by-example is often viewed as a natural mechanism that establishes certain patterns of behavior within the group, in particular because it shapes the beliefs of the followers. Gächter and Renner (2018) study the effect of leading-by-example in a repeated public goods game with groups of four people and find that leaders are 'role models', i.e. they shape followers' beliefs about the contributions of other players in their group, which then translate into their own contributions (see also Fischbacher and Gächter, 2010, for earlier accounts on the role of beliefs). This effect holds in all periods and is particularly strong in the first one; however, followers placed more weight on the observed contributions of other followers than those of the leaders in later periods.

I.2.2 Literature on Public Goods Games with Asymmetric Groups

Aside from coordination issues, differences among players prompt concerns about their perceptions of fairness. The literature documents the rise of normative conflicts due to the co-existence of multiple plausible contribution rules (for example, to contribute equally, efficiently, or make contributions that lead to equal payoffs). Such multiple rules can arise in groups whose members have different endowments, varying productivity or heterogeneous returns from the public good.

There are multiple ways in which heterogeneity among players is generated in public goods experiments. Some papers allow asymmetric payoffs by making players differ in their benefits from the public good. Van Dijk and Wilke (1993) and Fisher et al. (1995) were the first to find that asymmetry in returns from the public good among group members generates differences in contributions between players. Later works demonstrate that heterogeneity in marginal per capita returns (MPCR) can reduce contributions (see Fischbacher et al., 2014). Nikiforakis et al. (2012) find that when punishment is allowed, it is much more likely to trigger counter-punishment and start a feud in groups with asymmetric returns (that lead to normative conflicts) than in a setting with symmetric players. They also find that while the possibility of punishment sustains cooperation over the whole length of the experiment, the feud is very costly and fully offsets the efficiency gains from increased cooperation.

[6] However, if a "leader" does not serve as an example, like in Cox et al. (2010), this can result in lower earnings of the players.

Another way to generate asymmetry in payoffs is to allow for heterogeneity in players' endowments. Several studies report that heterogeneity in endowments reduces cooperation, especially when the distribution of asymmetric endowments is common knowledge (see Anderson et al., 2004; Cherry et al., 2005; Kroll et al., 2007; Oxoby and Spraggon, 2006), while others find no differences in the presence of asymmetric endowments (e.g. Buckley and Croson, 2006; Chan et al., 1996).

Some papers implement heterogeneity in both endowments and returns (see Chan et al., 1999, 2012), while others use heterogeneity in the productivity of contributions (Noussair and Tan, 2011; Tan, 2008). While all of the papers discussed above mainly focus on comparing asymmetric with symmetric groups, none of them explore the extent to which the composition of returns in a group affects the provision of the public good. To the best of my knowledge, I am the first to address this question.

Reuben and Riedl's (2013) paper on asymmetries and associated conflicting norms in a public goods game is closely related to my research question. In their paper, the authors study the emergence and enforcement of contribution norms in both symmetric and asymmetric groups of three people under various types of heterogeneity. They implement four types of groups: homogeneous with equal endowments, homogenous with unequal endowments, homogenous with unequal endowments and restricted contributions (with contributions restricted to the level of the smallest endowment in the group), and heterogeneous groups with unequal marginal benefits (but equal endowments). They conducted an online questionnaire study to elicit normatively appealing rules of behavior in the public goods game with individuals uninvolved in the experimental study. The analysis of the results of the questionnaire allowed the authors to select several contribution rules that appear to serve as normatively appealing rules of behavior. Their experimental study comprises eight treatments (each played for ten consecutive periods) that varied along two dimensions: the type of group and whether or not players could punish other group members (with a cost-to-punishment ratio equal to 1:3).

Reuben and Riedl document that with punishment, strong and stable differences in contributions emerge across groups that have either asymmetric endowments or heterogeneous MPCR. Here, I will focus on treatments with heterogeneous groups (unequal MPCR). When marginal benefits from the public good are unequal, the authors find a disagreement between players with high and low returns regarding the relative contribution rule to be enforced: low-benefiting players tend to enforce contributions that lead to equal payoffs, while high-benefiting ones enforce equal contributions. Although I do not allow for punishment in my experiment, possible disagreements regarding the relative contribution rule also play an important role in my setup.

I.2.3 Leadership in Asymmetric Groups

There are few papers on leading-by-example that study its effect in asymmetric groups. For example, Au and Chung (2007) and Neitzel and Sääksvuori (2013) vary the order of moves of heterogeneously-endowed players in a sequential game and find that setups where rich players move first generate higher contributions than alternative sequential move mechanisms (e.g. where poor players move first).

Perhaps the most closely-related paper on leadership to the present work is Levati et al. (2007), who study the effect of leading-by-example in a public goods game with asymmetrically-endowed individuals. Their 3×2 experimental design comprises six treatments with three different types of games (control, leading-by-example, and strong leadership with ostracism power) and two information conditions about the distribution of endowments (complete and incomplete). Each treatment lasts 24 periods in partner matching, with two group members endowed with 30 experimental currency units (ECU), and two members endowed with 20 ECU. In treatments with incomplete information, participants know their own endowment and that the total group endowment is 100 ECU, while in treatments with complete information the distribution of endowments is common knowledge. In leadership treatments, there is an exogenous part (periods 1 to 16) and an endogenous part (periods 17 to 24). In the exogenous part, each of the four players was appointed as the leader for four consecutive periods, with a predetermined and commonly-known sequence of taking turns. In strong leadership treatments, the leader can exclude one group member in the next period.

The authors find that leadership increases contributions in both symmetric and asymmetric groups, although this effect is only present in asymmetric groups when individuals are fully aware about the distribution of endowments. Moreover, it emerges that ostracism power does not increase cooperation further over leading-by-example in asymmetric groups compared with symmetric groups, but only in case of full information. If players are uncertain about the distribution of endowments, the effect is weaker for both leading-by-example and leadership with ostracism.

The key difference of the present study compared with Levati et al. (2007) is that I employ a different type of heterogeneity. This leads to a number of issues that pose tensions between allocations of contributions that cannot occur in setups with only asymmetric endowments. For example, this implies different relative contribution norms: to achieve equality in payoffs, the distribution of contributions under asymmetric endowments should differ from the distribution of contributions under asymmetric returns from the public good. Furthermore, with asymmetries in returns, some allocations of leaders' contributions restrict the number of possible contribution norms with which to comply, which cannot happen with only asymmetries in endowments (e.g. when low-benefiting players give up on equal payoffs to achieve higher efficiency, see detailed discussion in Section I.4). In addition, I only consider the "full information" case in which the distribution of returns is common knowledge, and I do not allow for exclusions of group members. Moreover, the appointment of leaders in Levati et al. (2007) allows for strategic cooperation in earlier periods and it can bias the results towards higher cooperation rates in leadership treatments, because the order of moves is predetermined and commonly known, while my approach excludes such a possibility.

I.3 The Experiment

The linear public goods game used in the experiment is designed as follows. There are n players in each group. Each player i is endowed with some number of points (game currency) y_i. Each player has to decide how much to contribute to the public good, c_i, and how much to keep for private consumption. Each point kept leads to a one-point increase in private earnings, while each point contributed to the public good (joint project) increases the earnings of all players. The players' payoffs, π_i, are determined according to the function

$$\pi_i = (y_i - c_i) + \gamma_i \sum_{j=1}^{n} c_j, \quad i = 1, \ldots, n, \tag{I.1}$$

where $\gamma_i \in \{\gamma_L, \gamma_H\}$ represents an individual's MPCR from the public good, with $\gamma_H = 0.75$ and $\gamma_L = 0.5$, y_i is i's endowment, and c_i is i's contribution.[7]

If all players were symmetric, they would possess equal endowments, $y_i = y$ for all i, and they would receive equal MPCR from the public good, $\gamma_i = \gamma$ for all i. In the setup that I employ, players only differ in their returns from the public good, γ_i. The higher the MPCR, the more that the agent benefits from the public good.[8]

In what follows, I use the same MPCR as in Reuben and Riedl (2013) and Kube et al. (2015), namely $\gamma_H = 0.75$ for the players with a high return, and $\gamma_L = 0.5$ for the players with a low return. I use groups of $n = 3$ players. There is always one follower with a high return ($\gamma_H = 0.75$) and the other follower with a low return ($\gamma_L = 0.5$). The main treatments vary in leaders' returns: high returns in the H-HL treatment and low ones in the L-HL treatment (these abbreviations reflect the composition of returns in the group, with the first letter separated by a hyphen standing for the leaders' return, and the two letters afterwards corresponding with followers' returns). I compare followers' responses to the contributions of leaders with different returns.

The heterogeneity in MPCR poses endogeneity problems if leaders' contributions differ between treatments. These endogeneity problems make it difficult to interpret and compare followers' reactions. To solve this, I need to isolate the effect of leading-by-example from leaders' considerations regarding the contribution. For this purpose, I use the strategy method (Selten, 1967) to measure followers' *cooperation preferences* (see Fischbacher and Gächter, 2010), namely their conditional contributions with respect to all possible contributions of the leader.[9] Subjects play a one-shot sequential public goods

[7]Note that the payoff function (II.1) represents a social dilemma if and only if $1/n < \gamma_i < 1$. The first inequality implies that the social optimum is for every subject to contribute all resources to the public good. The second inequality shows that individual returns from the public good are lower than private returns. Hence, if all agents are self-interested, the unique Nash equilibrium is to contribute zero.

[8]Alternatively, one could employ heterogeneity in endowments (Levati et al., 2007; Neitzel and Sääksvuori, 2013; Reuben and Riedl, 2013), and heterogeneity in productivity (Noussair and Tan, 2011; Tan, 2008), but final allocations under those types of heterogeneity exhibit no such tension between the equality of payoffs and the equality of contributions.

[9]In this conditional contribution a subject has to indicate (in an incentive-compatible way) how much she contributes to the public good given all possible contributions of the leader (see also Frackenpohl et al., 2016, for a similar approach). Participants were asked to enter their corresponding contribution for each

game: the leader makes contribution decisions first, then the two followers make their decisions by filling out a contribution table. The payoffs are computed using realized contributions according to the one stated by the leader and respective responses in followers' contribution tables.

In addition to cooperation preferences, I elicit players' beliefs about contributions of others:[10] leaders' beliefs on contributions of the followers (after the leader decided on contribution himself), and followers' beliefs on contributions of the other follower in the group, conditional on the leader's contribution.

In order to understand the behavior in a dynamic setting and compare the effect of leading-by-example with the simultaneous public goods game with heterogeneous agents, I further conduct additional (within-subject) treatments in which participants make contribution choices in a repeated linear public goods game.

After playing the one-shot game, participants were asked to play the repeated version of a public goods game. Players remained in the same groups of three people as before, but these groups were randomly allocated to two additional treatments: half of L-HL groups and half of H-HL groups played a repeated sequential public goods game, and the rest played a repeated simultaneous game. The participants kept the same return from the public good throughout the experiment (same as in the one-shot game) and they had the same role in sequential repeated games as in the one-shot game (i.e. leaders in the one-shot game were leaders in the sequential repeated game, etc.). The composition of returns was known to players.

The repeated game is played for ten rounds in partner matching mode.[11] In order to prevent strategic considerations regarding the decision in the main treatments I only report the outcomes of the one-shot game at the end of the experiment. The additional treatments allow comparing contributions and outcomes between simultaneous and sequential games, because it is not possible to properly elicit contribution preferences in a simultaneous game without turning it into a sequential one. In each round in a repeated game, players were informed about each player's contribution and earnings.

I.3.1 Procedures

The instructions (see Appendix I.A) explained to the participants the payment system in a public good problem and the differences between players. After the participants had read the instructions, they had to answer control questions to ensure their understanding of the payoff structure and the dilemma situation. The experiment only started after all

of 21 possibilities in a contribution table. The one-shot game was played only once, and the participants knew this. This allows to elicit subjects' contribution preferences, without confounding preferences with strategic considerations.

[10] I follow the terminology suggested by Fischbacher and Gächter (2010) and call players' expectations *players' beliefs*. These beliefs are *not* the beliefs in game-theoretical notation, i.e. they are not the probabilities players put on nodes within an information set in games with incomplete information.

[11] Of course, partner matching mode in a repeated game allows for the strategic behavior on the one hand, but on the other hand it allows excluding unobserved heterogeneity in beliefs and cooperative preferences that occurs if one uses stranger matching mode, and it allows for a more precise update of beliefs.

participants had answered all questions correctly. Each player was endowed with $y = 20$ points and allocated these points between investments in the public good and private account. Players had full information on the composition of returns in their group, as well as the number of players and their endowments. I use neutral wording in the instructions and during the experiment (e.g. players of low and with high returns are referred to as players of type A and of type B, respectively, with known MPCR).

At of the end of the experiment, participants were asked to complete a questionnaire that included questions on gender, age, field of study, self-assessment of risk preferences, trust and power distance, choice determinants and suggested contributions of other players in their group, as well as participants' assessment of fairness of hypothetical allocations, computed according to different normatively appealing contribution rules (see Section I.4 below).

All of the experiments were computerized using the software z-Tree (Fischbacher, 2007). The experiment took place at the BonnEconLab in July 2017. The participants were mostly students from various disciplines from the University of Bonn. All participants were recruited using HROOT recruitment software (Bock et al., 2014). I conducted four sessions (two for H-HL and two for L-HL groups) with 24 participants.[12] All 93 participants were randomly assigned into groups of three people. Each participant was randomly assigned the role (leader or follower), while the returns (high or low) were determined by the role and the treatment. I use random assignment of returns to players within treatments (cf. Nikiforakis et al., 2012). Accordingly, I can exclude possible confounds due to reputation or intentions effects (signal obtained from the observed "earned" return).[13]

On average, participants earned 14 EUR (slightly above 16 USD), including a show-up fee of 4 EUR.[14] In addition to their earnings from the public goods experiment, participants were paid up to 2 EUR for correctly guessing actual contributions of others.[15] Each session lasted roughly 80 minutes (including random allocation of the participants into cubicles, quiz, questionnaire and payments).

I.4 Behavioral Predictions

In this section, I first provide the intuition for the behavioral predictions for each treatment under different assumptions on the shape of players' utility functions that lead to different contribution norms using the parameters of the design. Detailed solutions are provided in Appendix I.B. Subsequently, I use the predicted contribution behavior to stress

[12] In one session I had 21 participants due to no-show.

[13] Some papers establish differences in fairness considerations based on the justification of asymmetries (see Cherry et al., 2005; Kroll et al., 2007; Van Dijk and Wilke, 1993).

[14] During the experiment subjects earned their payoffs in points, at an exchange rate of 1 point = 0.1 EUR for the points earned in a one-shot game, and at a rate of 1 point = 0.02 EUR for the points earned in repeated games.

[15] Each participant was paid 1 EUR for making a correct guess. Although followers had to make 21 guesses, there was only one relevant actual contribution of the other follower, whereas for the leader both actual contributions of the followers were relevant: in order to make these guesses incentive-compatible, I paid 1 EUR for each correct guess of the leader.

the differences between treatments. The analysis of implied contributions allows me to formulate some testable hypotheses.

I.4.1 Contribution (Social) Norms and Outcomes

Different models of social preferences predict different equilibria in a public goods game, most often transforming the game into a coordination problem with multiple equilibria. Some of these equilibria are not empirically observed, while some observed equilibria are predicted by various models. Rather than using a particular form of social preferences, one can study the behavior in a game with respect to social norms, namely "customary rules of behavior that coordinate our interactions with others" (Young, 2008, p. 647). These social norms are viewed as a coordination device, becoming contribution rules that players might or might not follow.[16]

According to Reuben and Riedl (2013), the contribution rules[17] in a public goods game with heterogeneous players vary along two dimensions. The first is related to the maximization of collective welfare and can be referred to as the efficiency rule, and the second is related to "fairness" and can be referred to as the relative contribution rule. Fairness considerations may differ depending on which principle to apply: equality or equity. Equality is usually thought of as equal outcomes regardless of individual characteristics, while equity is generally interpreted as the proportional dependence of outcomes on some characteristic such as an individual return from the public good.

Depending on what to apply the principle of equality to, one can formulate contribution rules that seem to be most attractive to players. Reuben and Riedl (2013) select several normatively appealing contribution rules that can potentially arise in groups with heterogeneous returns that do not necessarily achieve efficiency in the provision of the public good: the equality of contributions and the equality of payoffs. I apply their approach to my setup.[18] The predictions made below rely on the assumption that all players follow the same normative rule.

Table I.1 summarizes the behavioral predictions for players with social preferences under different contribution rules.[19] Here, the equality-of-payoffs rule represents equilibrium outcomes in the model with inequality averse players (Fehr and Schmidt, 1999), while the efficiency normative rule can appear if one assumes instead that players have

[16] Necessary conditions for a social norm to exist are, first, players must know that such a (normatively desirable) behavioral rule exists; second, such a rule is sustained by some mechanism, for example, by shared expectations about an appropriate solution to a given coordination problem (e.g. focal points, see Schelling, 1960). Other mechanisms of norm enforcement can rely on social disapproval or punishment, see Young (2008) for the discussion.

[17] Called normatively appealing rules of contribution behavior in their paper.

[18] Reuben and Riedl also report the existence of another contribution rule, which they call "proportionality to benefits". This rule predicts that players contribute somewhere in between the equality of payoffs and the equality of contributions, namely as it follows from the name, their contributions are proportional to their *returns* from the public good. However, this rule is difficult to justify to exist from a theoretical perspective, as it is always strongly dominated by other contribution rules.

[19] The setup used for heterogeneous returns in Reuben and Riedl (2013) (UMB treatment) is similar to L-HL group in my design, while the setup used in Kube et al. (2015) (HET treatments) is similar to H-HL group.

preferences for efficiency (e.g. Charness and Rabin, 2002). The equality-of-contributions rule may arise if one assumes that players have the utility function with a quadratic loss from non-compliance to preferred contribution norm (introduced in Cappelen et al., 2007). Conditions for existence and the calculation of equilibrium contributions for social preferences corresponding with a particular normative contribution rule are presented in Appendix I.B.

Table I.1: Followers' contributions and the efficiency of the public good under different normative rules, provided that the leader's contribution equals x

		Normatively appealing contribution rule[a]		
		Efficiency	Equality of	
Group[b]	Variable[c]		Contributions	Payoffs
H-HL	c_l	20	x	x
	c_H	20	x	x
	c_L	20	x	$0.4x$
	S	120	$6x$	$4.8x$
L-HL	c_l	20	x	x $(x \leq 10)$
	c_H	20	x	$2x$ $(x \leq 10)$
	c_L	20	x	x (≤ 10)
	S	105	$5.25x$	$7x$ $(x \leq 10)$

[a]Based on Table 2 in Reuben and Riedl (2013).
[b]Groups of size $n = 3$. First letter corresponds to the return of the leader (H for high and L for low), while the latter two are the returns of the followers.
[c]c_l corresponds to the contributions of the leader, c_H corresponds to the contributions of H-followers, c_L corresponds to the contributions of L-followers, and S denotes the total efficiency of the public good (sum of payoffs from the public good).

I.4.2 Implications for Expected Players' Contributions

First, the calculations presented in Table I.1 suggest that followers' contributions must be higher in L-HL group than in H-HL group, given the same level of the leader's contribution. For illustration, it is useful to compare their responses under different contribution rules.[20] The equality-of-contributions rule suggests that followers' contribution decisions (with respect to the leader's contribution) should lie on a 45-degree line. The equality-of-payoffs rule implies that contributions of followers differ. Under this rule, H-followers in H-HL groups and L-followers in L-HL groups should contribute as much as the leader does. For L-followers in H-HL groups, conditional contributions should lie below the 45-degree line. For H-followers in L-HL groups, their conditional contributions should lie above the 45-degree line. Since some followers might prefer the equality-of-payoffs rule, their contributions will drive the difference in average conditional contributions.

[20]The efficiency rule clearly predicts that conditional contributions are constant at the value of 20. Note that this and the equality-of-contributions rules cannot be distinguished if the leader's contribution equals 20.

Second, it should be mentioned that the contribution of a leader as a coordination device being effective hinges upon the clarity of the "signal". Coordination on a certain contribution norm in different groups clearly depends on the contribution of the leader with a particular return. In L-HL groups, the leader's contribution should not exceed 10 under the equality of payoffs, otherwise the rule would be violated, since the follower with a high return cannot contribute more than his endowment. The cases when these limitations are violated should clearly indicate the leader's preferences for normative rules. For example, if the L-leader contributes more than 14, it clearly signals that her preferred contribution rule is the equality of contributions, and the preferred contribution rule is the efficiency if she contributes the full endowment. Note that there are no such limitations in H-HL groups,[21] and thus one can expect the "signal" obtained from the H-leader to be less clear.

Third, one may expect different contribution rules to be observed in groups led by leaders with different returns. One can show that if players are allowed to select between multiple contribution rules (see Cappelen et al., 2007), they should choose a particular rule depending on their own return from the public good.[22] In this case, players with high returns would prefer the equality-of-contributions rule, and players with low returns would prefer the equality-of-payoffs rule.

Fourth, to better understand the differences in potential outcomes under different normative rules, consider the following thought experiment: if every player in the group follows the equality-of-payoffs rule, then the maximum possible contributions that adhere to this rule must be (10,20,10) for the L-HL group and (20,20,8) for the H-HL group (the first number is the contribution of the leader, the second is the contribution of the H-follower, and the third is that of the L-follower). This leads to equal payoffs of 30 (36) points for each player in the L-HL (H-HL) group. Under the efficiency rule (with all players contributing 20 to the public good), the payoffs of corresponding players are (30, 45, 30) in the L-HL and (45,45, 30) in the H-HL groups, respectively. Individual gains (or costs) from switching to the efficiency rule from the equality-of-payoffs rule are (0,15,0) for L-HL and (9,9,-6) for H-HL groups, correspondingly. For L-HL groups, it leads to an improvement in the efficiency with no monetary costs for players, while increased payoffs of H-players in H-HL groups come at the cost of the L-player's payoff. Whether L-players in L-HL groups give up on inequality to achieve higher efficiency is an empirical question.

Finally, H-players in H-HL groups receive higher payoffs if they both contribute positive amounts than if they do not contribute at all (i.e. assuming that the H-follower does not free-ride) even if the L-follower completely free-rides. Therefore, it is another interesting empirical question whether players with high returns in H-HL groups cooperate regardless of whether the L-follower follows or not.

[21] The H-leader's contribution does not help to distinguish whether the leader suggests the equality-of-payoffs rule or the equality-of-contributions rule.

[22] In a similar way it is done in Neitzel and Sääksvuori (2013) for heterogeneously-endowed players. I do this exercise in Appendix I.B.

I.5 Results

In this section, I report the results from the analysis of contributions from the main treatments in my experiment, i.e. from the one-shot games.[23] After presenting summary statistics, I start with non-parametric tests of realized contributions and payoffs to ascertain whether there are any differences in contribution decisions and associated payoffs. Subsequently, I take a closer look at the contribution and belief profiles of the followers. Finally, I estimate a random effects model for followers' conditional contributions.

My findings suggest that, first of all, the contributions of players differ depending on the composition of returns (the leader's return and the follower's own return). In particular, I find that followers contribute more in groups with L-leaders than in groups with H-leaders, given the same level of the leader's contribution. Moreover, L- and H-followers contribute closely in groups with L-leaders, while in groups with H-leaders followers' contributions differ depending on their returns. There are also differences in expectations about the contributions of the other follower in the group, and how followers with different returns react to those expectations: while L-followers contribute less than they expect the H-follower in their group to contribute, the H-followers contribute as much as they expect the L-follower in their group to contribute. Finally, regression analysis shows that individual contribution preferences are systematically affected by the composition of returns, whereby the differences in conditional contributions remain present even after I control for heterogeneity in beliefs and individual characteristics.

I.5.1 Contributions and Payoffs in the One-Shot Game

Table II.2 shows summary statistics on contributions and payoffs (in points) of players with different returns by group. It appears that leaders' average contributions do not differ, but followers' average actual contributions vary quite a lot. The largest difference is observed in contributions of L-followers between treatments, whereby their contributions are the lowest in H-HL groups. However, median contributions show that contributions of leaders are in fact unequal, and that all players contribute more generously in the L-HL treatment.

In order to compare whether actual contributions and actual payoffs statistically differ between followers with different returns and whether those differ between leaders and followers, I apply a series of Wilcoxon signed-rank tests to the realized values of contributions and payoffs. The tests show that realized contributions of H-followers differ from those of L-followers in H-HL groups ($p = 0.035$), but not in L-HL groups ($p = 0.92$).[24] At the same time, the payoffs realized by the followers are nearly identical in H-HL groups ($p = 0.58$) but differ in L-HL groups ($p = 0.023$).

[23] Computations for the analysis presented in this paper were programmed in Python (except for the regression analysis) using the Templates for Reproducible Research Projects in Economics by von Gaudecker (2014).

[24] All p-values are for two-sided tests, if not mentioned otherwise.

Table I.2: Summary statistics for the **realized** outcomes in the one-shot game (exogenous returns)

Player	# obs	Mean	St.dev.	Min	Max	Median
Contributions						
H-HL, leader	15	10.9	7.34	0	20	10
H-HL, L-follower	15	3	5.92	0	20	0
H-HL, H-follower	15	8	9.11	0	20	1
L-HL, leader	16	10.9	8.97	0	20	13.5
L-HL, L-follower	16	9	8.85	0	20	6.5
L-HL, H-follower	16	9.31	9.08	0	20	8.5
Payoffs						
H-HL, leader	15	25.5	9.28	15	42.5	20
H-HL, L-follower	15	27.9	6.35	20	40	27
H-HL, H-follower	15	28.4	6.54	20	42.5	28.8
L-HL, leader	16	23.7	4.28	20	32.5	22.5
L-HL, L-follower	16	25.6	6.32	17.5	37.5	24.2
L-HL, H-follower	16	32.6	11.3	20	50	33.8

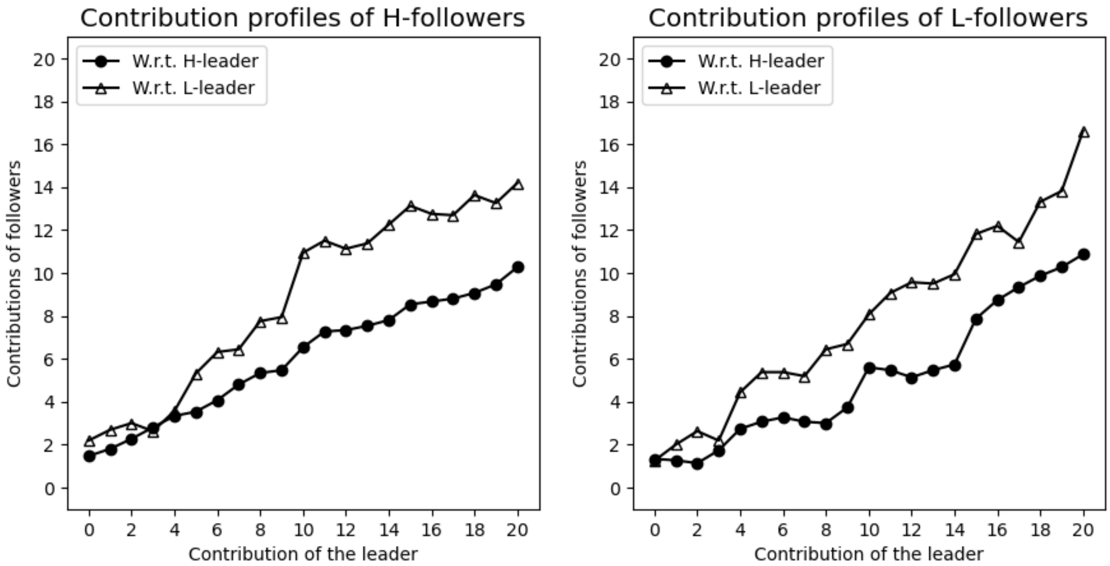
Regarding the differences between followers and leaders, L-followers contribute on average significantly less than H-leaders in H-HL groups (Wilcoxon signed-rank test $p = 0.007$), although tests show no other differences between followers and leaders ($p > 0.26$). H-followers in L-HL groups receive much higher payoffs than leaders ($p = 0.003$), but no test shows differences between the payoffs of leaders and followers in any other cases ($p > 0.26$).

One way to interpret these observations is that followers might adhere to different contribution rules depending on their own return: H-followers seem to prefer equal contributions, while L-followers seem to prefer equal payoffs. This can also be viewed differently: players in L-HL groups seem to follow the equality-of-contributions rule, while players in H-HL groups seem to follow the equality-of-payoffs rule. All of these results align well with the previously-described behavioral predictions, although it is difficult to test whether a particular follower adheres to a particular contribution rule.

I.5.1.1 Conditional Contributions and Conditional Beliefs

The differences in realized contributions described above seem to stem from heterogeneous followers' responses because average leaders' contributions are equal. Figure I.1 shows averaged contribution profiles (conditional contributions) of the followers with respect to leaders with different returns. Although Figure I.1 shows that followers' responses differ depending on whether the leader has high or low returns, Figure I.2 shows that these differences are driven by the leader's return from the public good.

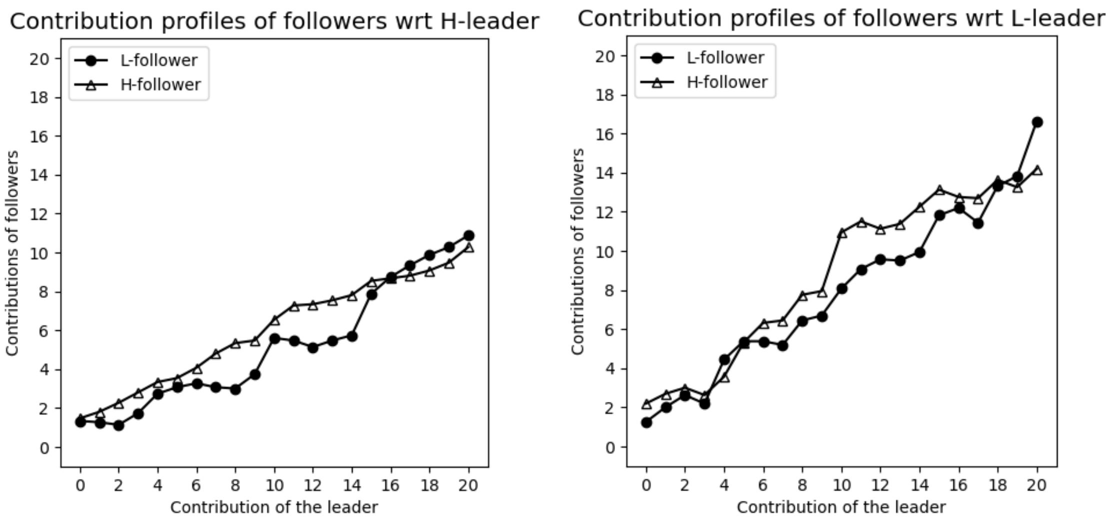
Since the contribution profiles presented above are *average* conditional contributions of all followers with a particular return, comparing them directly might miss underly-



(a) H-followers

(b) L-followers

Figure I.1: Averaged contribution profiles of followers with respect to different leaders (exogenous returns)

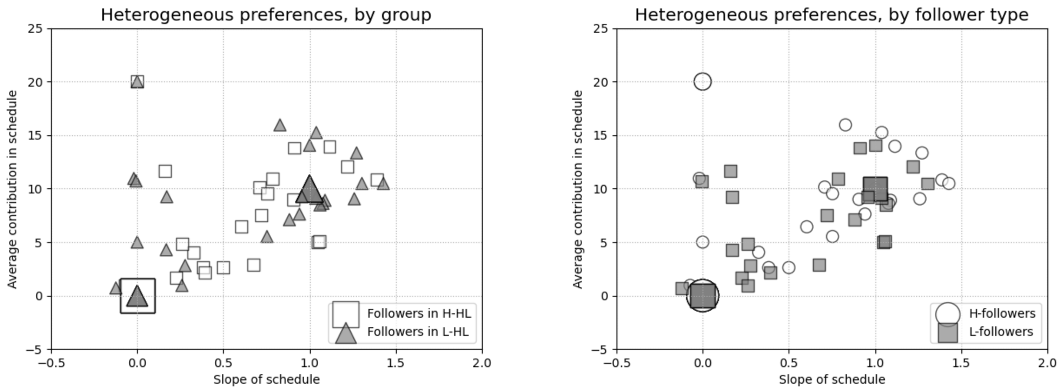


(a) H-HL groups

(b) L-HL groups

Figure I.2: Averaged contribution profiles of followers with respect to identical leaders (exogenous returns)

ing heterogeneity. Below, I compare individual profiles of the followers by contrasting the slopes estimated from a linear regression on the leader's contribution and average contributions of their contribution profiles in the same way as in Fischbacher and Gächter (2010). Scatter plots in Figure I.3 depict individual contribution preferences of the followers, with the mean contribution in the schedule (on the y axis) versus the slope of the individual schedule with respect to the leader's contribution (on the x axis). These scatter plots suggest that the differences described above did not occur by chance due to specific contribution preferences of the followers in H-HL and L-HL groups, but rather they are systematic. The scatter plot in Figure I.3a suggests that contribution preferences of followers in L-HL groups differ from those of followers in H-HL groups, as the former appear to form two clusters (around estimated slopes equal to zero and one, correspondingly), while the latter spread in between. Mann-Whitney U tests confirm the difference: $p = 0.049$ for slopes and $p = 0.024$ for means of the schedules.[25] At the same time, there seem to be no differences between contribution preferences between H- and L-followers in general, as plotted in Figure I.3b (Mann-Whitney U tests $p > 0.38$).



(a) By group (treatment)

(b) By follower

Figure I.3: Heterogeneous contribution preferences (exogenous returns)

Figures I.1, I.2 and I.3 point to several important observations. First, followers are on average (imperfect) conditional cooperators. [26] Second, followers seem to follow L-leaders quite closely, at least more closely than H-leaders. These findings can be summarized in the following result.

Result 1. *For a given contribution of the leader, both H- and L-followers contribute more in L-HL groups than in H-HL groups.*

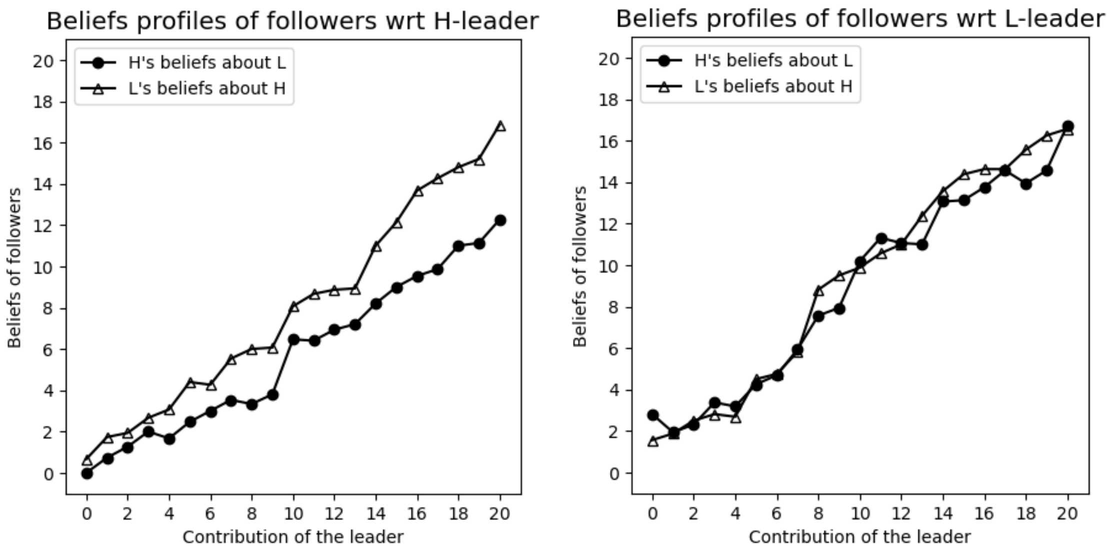
Figure I.4 provides more support for the hypothesis that players may select the normative contribution rule taking into account the group composition. Figure I.4 shows averaged beliefs of the followers about the contribution of the other follower in their group conditional on the contribution of the leader.[27] L- and H-followers in H-HL groups show different expectations about contributions of the other follower in their group, while no substantial differences are observed in L-HL groups. The important observations here are

[25] According to the classification used in Fischbacher and Gächter (2010), perfect conditional cooperators should have mean contributions of 10 and the slope of 1, free-riders should have zero mean contributions and zero slope, and unconditional cooperators should have zero slope, but positive mean contributions.

[26] See also individual contribution patterns with respect to the leader's contributions, presented in Appendix I.C.

[27] Followers' conditional beliefs are also presented in Appendix I.C. I do not denote followers' ids on these graphs, but subplots for conditional contributions and conditional beliefs in a particular position correspond with the same follower, i.e. top left subplot for contributions and top left subplot for beliefs are plotted for the same individual.

that, first, followers also expect other followers to be (imperfect) conditional cooperators, and second, L- and H-followers in H-HL groups show different expectations about the contributions of the other follower in their group, while there seems to be no difference in L-HL groups.



(a) H-HL groups

(b) L-HL groups

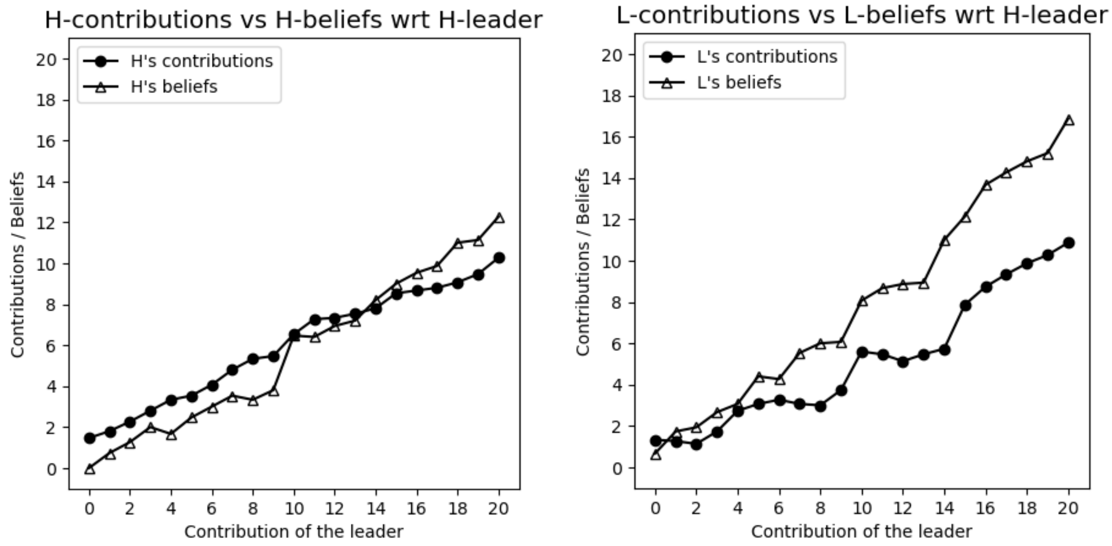
Figure I.4: Averaged beliefs profiles of followers with respect to leaders' contribution decisions (exogenous returns)

How are these beliefs reflected in followers' own contributions? Figures I.5 and I.6 show how followers' conditional contributions are related to their beliefs about contributions of the other follower (conditional on the leader's contribution). Conditional contributions of L-followers differ from their beliefs quite substantially in H-HL groups but only slightly in L-HL groups. For H-followers, there seems to be no such distinction.

Regarding leaders' beliefs, both H- and L-leaders expect L-followers to contribute slightly less than leaders themselves, although they expect H-followers to match leaders' contributions on average. However, these differences are not statistically significant.

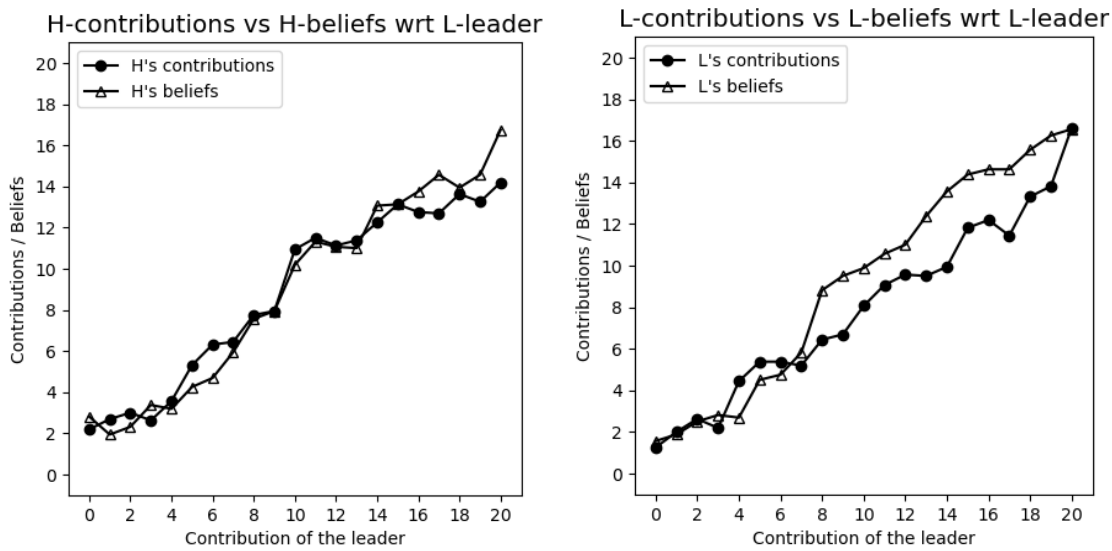
Result 2. *In both H-HL and L-HL groups, L-followers contribute less than they expect H-followers to contribute, but H-followers contribute as much as they expect L-followers to contribute.*

This result is in line with Reuben and Riedl (2013), who find that players with low returns tend to enforce equal payoffs while players with high returns tend to enforce equal contributions. In my case, this is reflected in the patterns of conditional contributions and conditional beliefs: L-followers contribute on average less than they expect H-followers to contribute, and H-followers contribute on average approximately the same as they expect L-followers to contribute. It also seems that three factors are important for conditional contributions: the leader's return from the public good, the leader's contribution (con-



(a) H-followers (b) L-followers

Figure I.5: Averaged contribution profiles and own beliefs about the other follower in **H-HL groups** (exogenous returns)



(a) H-followers (b) L-followers

Figure I.6: Averaged contribution profiles and beliefs in **L-HL groups** (exogenous returns)

ditional on a given leader's return), and followers' own beliefs about the contribution of the other follower. Regression analyses below test these relations in a more rigorous way.

I.5.2 Regression Analysis

In order to analyze in a systematic way how the composition of returns in the group, leaders' contributions, and beliefs (about the other follower's conditional contribution) transform into followers' conditional contributions in the one-shot game, I estimate the following econometric model:

$$
\begin{aligned}
C_{i,j} = \alpha_0 + \beta C_l(j) + \gamma\ B_{i,j} \\
&+ \alpha_{leader}\ \textit{H-leader} + \alpha_{own}\ L\text{-}follower + \alpha_{HL}\ \textit{H-leader} \times L\text{-}follower \\
&+ \beta_1\ \textit{H-leader} \times C_l(j) + \beta_2\ L\text{-}follower \times C_l(j) \\
&+ \beta_3\ \textit{H-leader} \times L\text{-}follower \times C_l(j) \\
&+ \gamma_1\ \textit{H-leader} \times B_{i,j} + \gamma_2\ L\text{-}follower \times B_{i,j} + \gamma_3\ \textit{H-leader} \times L\text{-}follower \times B_{i,j} \\
&+ \alpha_X\ X_i + \varepsilon_{i,j}, \qquad \text{(I.2)}
\end{aligned}
$$

where $C_{i,j}$ is the dependent variable, the conditional contribution in response to the leader's contribution, $C_l(j) \equiv j$. $B_{i,j}$ is the own belief about the contribution of the other follower in response to the leader's contribution, $L\text{-}follower$ is a dummy equal to 1 if the follower has a low return from the public good (type A in the instructions) and 0 otherwise, *H-leader* is a treatment variable, namely a dummy equal to 1 if the leader has a high return (1 for H-HL groups) and 0 otherwise, X_i is the set of individual characteristics, and $\varepsilon_{i,j}$ is a random effects error term.

Table I.3 presents the pooled sample estimates for three econometric specifications for a random effects model for the conditional contributions of the followers.[28] Column 1 presents the simplest specification, which is a RE model for own (conditional) contribution with respect to the leader's contribution, dummies for composition of returns in the group, and the set of interaction terms to capture systematic differences in patterns. The model presented in column 2 adds the follower's own beliefs about the (conditional) contributions of the other follower in the group, along with corresponding interaction terms capturing potential heterogeneity in beliefs. Finally, the model presented in column 3 also includes additional measures for individual characteristics from the questionnaire, namely gender, self-assessed risk preferences, trust, and power distance.[29] All standard errors are clustered at the individual (participant's) level.

The results presented in Table I.3 show that the leader's contribution is almost as important for followers' contribution decisions as the beliefs about the other follower's contributions. However, in my specifications (unlike in Fischbacher and Gächter, 2010; Gächter and Renner, 2018) the sum of the coefficients for the leader's contribution and own beliefs is statistically different from unity (Wald tests $p = 0.074$ in model 2 and $p = 0.080$ in model 3). On average, in groups led by H-leaders, followers contribute less with respect to the leader's contribution than in groups with L-leaders, but no other statistically significant differences were found in the pooled regression.

Result 3. *In their contribution decisions, followers take into account both their beliefs*

[28] Randomization check (Hotelling T-squared generalized means test) shows that the differences in observables (contributions, beliefs) between treatments are not random ($p = 0.000$).

[29] Additional variables were transformed into dummies reflecting gender, high risk aversion, high risk loving, very high/low trust, and strong/weak power distance. These additional variables are not reported in the regressions below.

Table I.3: Estimates for a random effects model of conditional contributions (exogenous returns)

	(1) Contribution	(2) Contribution	(3) Contribution
Leader's contribution	0.667***	0.323**	0.320**
	(0.137)	(0.124)	(0.124)
L-follower	-1.043	-0.706	-1.141
	(1.790)	(1.737)	(2.104)
H-leader	-0.523	0.271	-0.004
	(2.058)	(2.182)	(2.327)
L-follower × Leader's contribution	0.020	0.142	0.142
	(0.178)	(0.174)	(0.174)
L-follower × H-leader	-0.349	-0.956	-1.193
	(2.453)	(2.383)	(2.460)
H-leader × Leader's contribution	-0.226	-0.235	-0.232
	(0.180)	(0.133)	(0.132)
L-follower × H-leader × Leader's contribution	0.036	-0.273	-0.271
	(0.240)	(0.197)	(0.196)
Belief		0.449***	0.452***
		(0.124)	(0.124)
H-leader × Belief		0.130	0.124
		(0.199)	(0.198)
L-follower × Belief		-0.185	-0.185
		(0.211)	(0.211)
L-follower × H-leader × Belief		0.283	0.281
		(0.283)	(0.284)
Constant	2.122	1.565	-0.186
	(1.551)	(1.621)	(2.354)
Other controls	No	No	Yes
Observations	1302	1302	1302
Number of individuals	62	62	62

Notes: *L-follower* is a dummy equal to 1 if the follower has a low return from the public good; *H-leader* is a dummy equal to 1 if the leader in the group has a high return from the public good (1 for H-HL groups); *Belief* is a belief from a conditional beliefs table about how much the other follower in the group contributes in response to a particular contribution by the leader; additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

about the other follower's contribution and the contribution made by the leader.

Although the results in Table I.3 show some differences in the contribution preferences of followers, there might be important interplays between the observed individual characteristics and treatment variable. In a series of robustness checks, I estimate regressions in (II.1) using the sub-samples of followers in groups with H- and L-leaders separately. The results are presented in Table I.D.1 in Appendix I.D. Additionally, Table I.D.2 in Appendix I.D presents the estimates of the same model separately on the sub-samples of L- and H-followers. The results reported in these tables confirm that followers respond differently to the contributions made by leaders with varying returns depending on their own and the leader's returns.[30]

Result 4. *Individual contribution preferences are systematically affected by the composition of returns in the group, whereby regression analysis shows that the differences in conditional contributions remain present after individual characteristics and heterogeneity in beliefs are taken into account.*

It emerges that followers are more responsive to the contribution of L-leaders.[31] In order to test whether followers' responses to leaders' contributions differ depending on the leader being "alike", I estimate the random effects model for followers' conditional contributions with respect to the contribution of the leader who has the same or the opposite return from the public good as the follower in consideration. Table I.4 below shows the estimation results. The variables "contribution by leader with same return" and "contribution by leader with different return" equal the leader's contribution if the leader and the follower have the same/different returns, and zero otherwise, and vice versa. The variables, reflecting followers' beliefs about the contribution of the other follower were defined in a similar way. These results clearly show that L-followers only respond to the contribution by L-leaders, while H-followers respond to leaders with both returns, although they follow L-leaders more closely.[32]

Result 5. *Followers follow L-leaders more closely than H-leaders, ceteris paribus, but L-followers follow L-leaders only.*

The results reported in this section suggest that my main theoretical predictions hold true: followers are more responsive to L-leaders, and they contribute more in L-HL groups than in H-HL groups in response to the leader's contribution. Differences in contribution patterns imply that players' contribution preferences are affected by group composition, possibly reflecting different contribution norms triggered by contributions of the leaders with varying returns. Although these differences are not very pronounced in the pooled sample, they become very salient in the sub-samples of followers with a particular return.

I.6 Repeated Games

I.6.1 Motivation

Do players make decisions solely according to their contribution preferences or are their decisions driven by other possible motives when it comes to a repeated interaction? How are players' actions in a sequential game related to their actions in a simultaneous game,

[30] Tables I.D.3, I.D.4, and I.D.5 in Appendix I.D also present estimates for a Fixed Effects model of conditional contributions, similar to (II.1). These tables show the robustness of the results reported above.

[31] Interaction terms "H-leader x Leader's contribution" in Table I.3 and Table I.D.2 in Appendix I.D, as well as the coefficient for "Leader's contribution" in Table I.D.1 in Appendix I.D provide support for this hypothesis.

[32] Fixed effects model estimates, reported in Table I.D.6 in Appendix I.D confirm the robustness of this result.

Table I.4: Estimates for random effects model of conditional contributions with respect to contributions by the leader with similar/different (exogenous) returns

	(1) H-followers	(2) L-followers
Leader with same return	-0.778	1.766
	(2.132)	(1.257)
Contribution by leader with same return	0.085	0.457***
	(0.047)	(0.123)
Contribution by leader with different return	0.320*	-0.052
	(0.125)	(0.079)
Belief, when leader has same return	0.582***	0.273
	(0.156)	(0.171)
Belief, when leader has different return	0.452***	0.687***
	(0.125)	(0.107)
Constant	-0.419	0.302
	(2.549)	(2.192)
Other controls	Yes	Yes
Observations	651	651
Number of individuals	31	31

Notes: *Belief* is a belief from a conditional beliefs table about how much the other follower in the group contributes in response to a particular contribution by the leader. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

in which no leader contributes first?[33] In order to answer these questions, one needs to understand how contribution behavior in a one-shot game translates into dynamics in repeated interaction.[34]

In a repeated game, players can strategically decide to contribute in earlier periods to build reputation for cooperation and thus receive higher payoffs at the end of the game. Consequently, in a repeated game it is more likely to expect higher contributions in earlier periods (compared with a one-shot game), which decline over time towards the end of the experiment. Moreover, in a repeated game players receive feedback on how members of their group performed in each round, which can confirm/update shared expectations about the behavior of others, which are necessary to sustain a certain contribution norm.

I.6.2 Results for the Repeated Games

In this subsection, I focus on how the contribution preferences elicited in the one-shot game translate into the overall dynamics of contributions when it comes to a repeated interaction. I analyze the differences between one-shot and repeated sequential games as

[33] The comparison of behavior in a sequential one-shot game with behavior in a simultaneous one-shot game seems problematic, since in a simultaneous game only actual contribution choices can be properly measured.

[34] Unfortunately, one can only compare actual contributions in groups with and without leaders to establish the effect of leadership, because conditional contributions imply some sort of sequentiality in decision-making, and cannot be properly elicited in a simultaneous game.

well as the differences between repeated sequential and repeated simultaneous games.

In repeated games, players can have strategic incentives to collaborate in the beginning, e.g. for building reputation for cooperation. Moreover, players in repeated games are able to update their beliefs about the contributions of other players according to the outcomes in previous periods. Nevertheless, my experimental design assures the absence of updating in the first round of repeated games, as the outcomes from the one-shot interaction are only reported at the end of the experiment. I start with the comparison of average contributions in the one-shot game versus the repeated sequential game to check whether strategic effects are present in my setup. Subsequently, I compare contribution dynamics in repeated sequential and repeated simultaneous games to ascertain whether leading-by-example helps to improve cooperation in the presence of heterogeneity in returns. Finally, I analyze the determinants of contribution decisions in the repeated games using regression analysis.

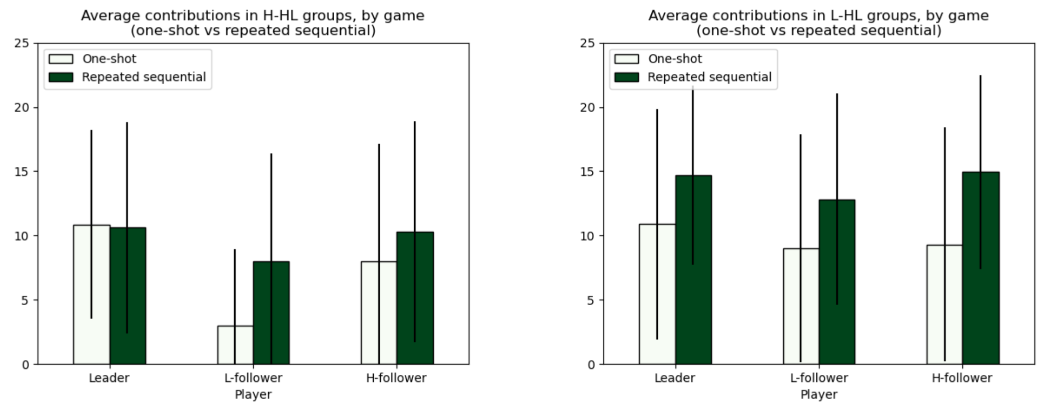
I find that followers demonstrate a strategic increase in contributions at the beginning of repeated games, compared with the contribution preferences elicited in the one-shot game. This effect is particularly strong in H-HL groups. Moreover, it is likely that the strategic incentives outweigh the effect of leading-by-example, as I do not find differences between average contributions in repeated sequential and repeated simultaneous games. Finally, the regression analysis of contributions in the repeated sequential game suggests that leaders try to act as role models during the whole course of the repeated sequential game, as their contributions do not decline over time.

Tables I.E.1 and I.E.2 in Appendix I.E show basic summary statistics (mean, standard deviation, and median) on actual contributions and payoffs in repeated games, separately for periods 1, 2–5, 6–9, 10, and over the whole game. Outcomes in repeated games differ from those in the one-shot game, as in both repeated sequential and repeated simultaneous games average contributions are much higher than realized contributions in the one-shot game. This increase is particularly strong for the first round of repeated games. Below, I try to ascertain what drives these differences.

I.6.2.1 One-Shot Versus Repeated Sequential Games

First, I compare contribution preferences elicited in the one-shot game with actual contribution decisions in the first round of the repeated sequential game. The reason for this comparison is that players have similar information sets in the one-shot game and the first round of the repeated sequential game, as they do not yet know the outcomes for the one-shot game. Figure I.7 shows average (actual) contributions of players in different groups. It appears that the only difference in H-HL groups is the higher average contributions of L-followers (Mann-Whitney one-sided U-test $p = 0.008$), while other players contribute on average the same ($p > 0.155$). At the same time, all players in L-HL groups contribute on average less in the one-shot game than in the repeated sequential one (Mann-Whitney one-sided U tests $p < 0.046$).

The observed differences can be driven by either different contributions of the leaders or



(a) H-HL groups (b) L-HL groups

Figure I.7: Comparison of average contributions of players between one-shot and repeated games (exogenous returns)

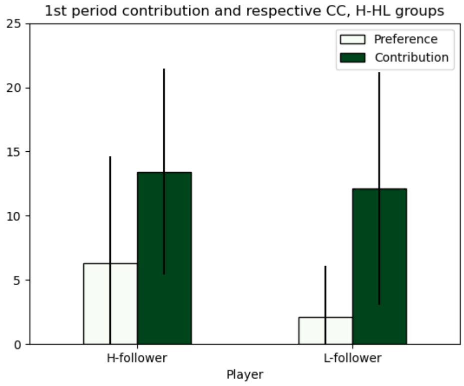
inconsistent reactions of followers. Figure I.8 shows that the latter is most likely the case: at least followers in H-HL groups contributed much more than they stated in their conditional contributions in the one-shot game. Graphs on the left (panels I.8a and I.8c) show average actual contributions of the followers in the first period of the repeated sequential game and their respective contribution extracted from the conditional contribution, given the contribution of the leader in the first period of the repeated game. Graphs on the right (panels I.8b and I.8d) show average deviations from the stated conditional contribution that appeared in the first period of the repeated sequential game.

Result 6. *Followers in H-HL groups seem to strategically increase their contributions in the first period of the repeated sequential game compared with the contribution preferences that they reported in the one-shot game. By contrast, followers in L-HL groups do not show much of a difference.*

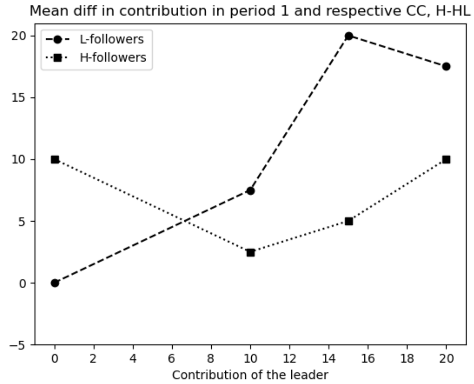
I.6.2.2 Repeated Sequential Games Versus Repeated Simultaneous Games

The second question that I aim to answer is whether the leading-by-example helps to mitigate social dilemmas in groups with heterogeneous returns when it comes to a repeated interaction. The results that I present below show that this is not necessarily the case: players demonstrate similar dynamics of contributions irrespective of whether they have a leader who contributes first and makes his contribution public or not. This effect is most likely driven by a strategic increase in contributions at the beginning of the repeated interaction, as we have already seen before in repeated sequential games.

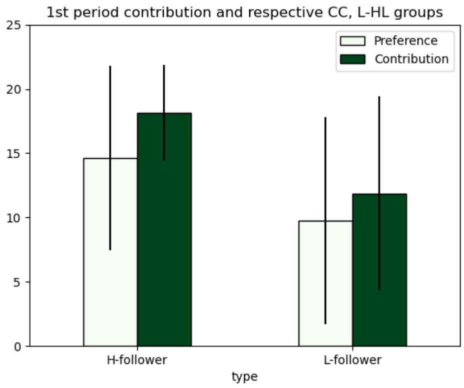
Figure I.9 compares average realized contributions of H- and L-players between repeated sequential and repeated simultaneous games. Mann-Whitney one-sided U tests show no statistical differences in contributions between two different types of repeated games ($p > 0.16$), except for L-followers in HHL groups ($p = 0.070$).



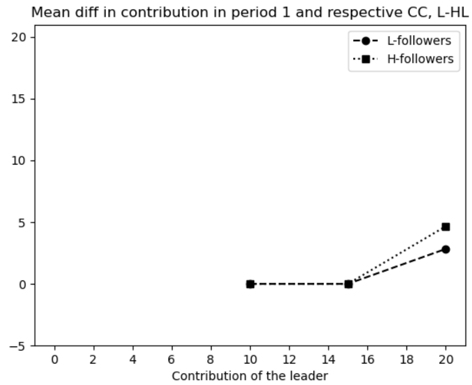
(a) Average contributions, H-HL groups



(b) Average differences, H-HL groups



(c) Average contributions, L-HL groups



(d) Average differences, L-HL groups

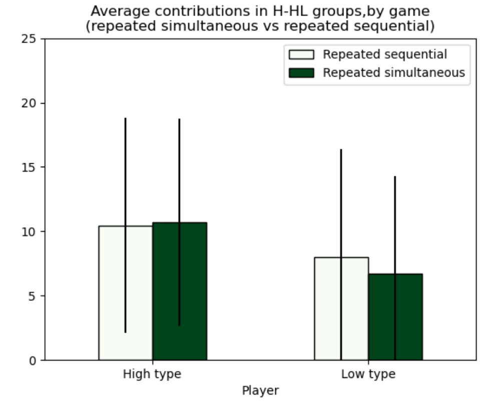
Figure I.8: Comparison of average contribution in the first period of repeated sequential game and corresponding contribution from the schedule (exogenous returns)

Result 7. *Contributions in a repeated simultaneous game do not statistically differ from contributions in a repeated sequential game, holding the composition of returns in the group constant. The only exception is L-followers in HHL groups, who contribute slightly more in a sequential game than in a simultaneous one.*

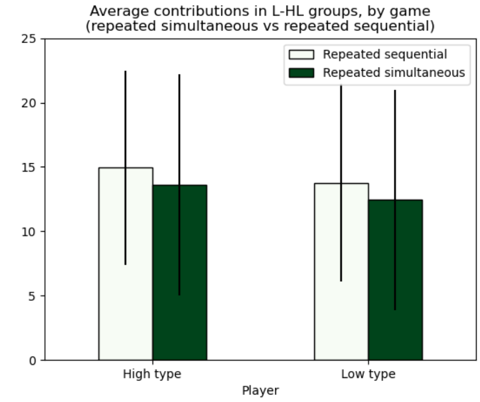
One potential explanation for this result is that strategic incentives for players are very strong and outweigh the effect of leading-by-example (this is supported by a similar result in Yamakawa et al., 2016).[35]

Figure I.10 shows the dynamics of average contributions in different repeated games and groups. All contributions decline towards the last period, although average initial

[35] Unfortunately, one cannot test whether players in groups with leaders better coordinate on a particular contribution rule, because multiple rules coexist, and one cannot distinguish whose contribution players use as a reference (or on which they put more weight when making their own contribution decisions). Nevertheless, based on contribution dynamics in repeated games, it is possible to assume a particular contribution rule to be followed in certain groups. Appendix I.F additionally provides graphs on individual dynamics of contributions. Some groups clearly agree on equal contributions in a sequential game (e.g. H-HL groups 2 and 6 and all L-HL groups except 4 and 7, see Figures C9 and C10) while some seem to prefer the equality of payoffs (e.g. H-HL group 4 and L-HL group 7), whereas in a simultaneous game, with rare exceptions, no such patterns occur.



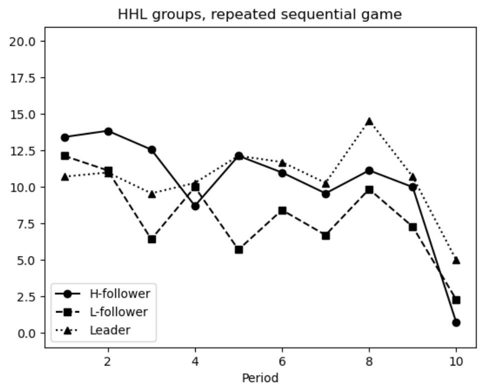
(a) HHL groups



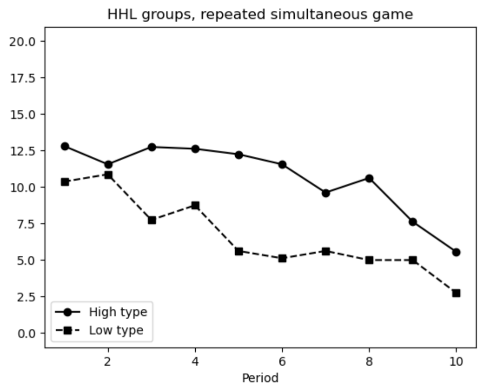
(b) LHL groups

Figure I.9: Comparison of average contributions of players between repeated simultaneous and repeated sequential games (exogenous returns)

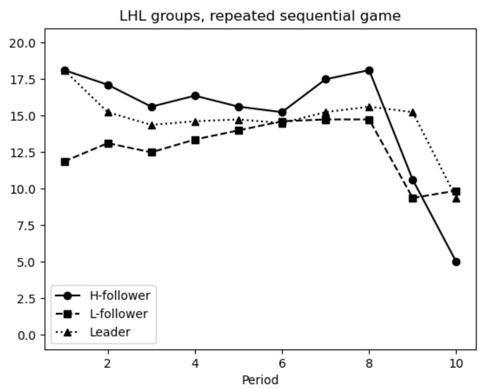
contributions in LHL groups are quite high, which results in better overall cooperation over the course of a repeated game.



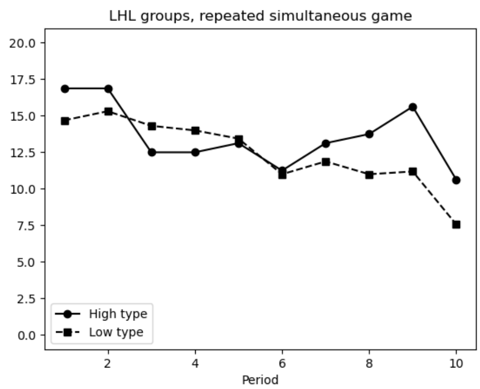
(a) HHL groups, sequential game



(b) HHL groups, simultaneous game



(c) LHL groups, sequential game



(d) LHL groups, simultaneous game

Figure I.10: Dynamics of average contributions in different types of repeated games (exogenous returns)

I.6.3 Regression Analysis

Below, I analyze what drives players' contributions in repeated games. Since in repeated games I do not elicit players' beliefs about contributions of others, I am unable to estimate regressions on belief formation. This limitation does not allow fully capturing the role of beliefs in the dynamics of contributions, especially in a repeated simultaneous game (in a repeated sequential game, the leader's contribution is a good approximation for initial beliefs, according to Gächter and Renner, 2018). Instead, I use the values extracted from the conditional beliefs schedule from the one-shot game with a corresponding leader's contribution. These imputed beliefs should most closely correspond with the beliefs in the first period of the repeated game (which, according to Gächter and Renner, 2018, are most strongly affected by the leader's contribution), but they do not necessarily coincide because players can anticipate a strategic increase in the contributions of others.

I estimate the following dynamic panel regression for the contributions of followers in a repeated sequential game:

$$\begin{aligned}C_{i,t} = \alpha_0 + \alpha_{leader} L_t + \alpha_{belief} B_{i,t} \\ + \alpha_{leader}\ H\text{-}leader + \alpha_{own}\ L\text{-}follower + \alpha_{HL}\ H\text{-}leader \times L\text{-}follower \\ + \beta_1\ H\text{-}leader \times L_t + \beta_2\ L\text{-}follower \times L_t + \beta_3\ H\text{-}leader \times L\text{-}follower \times L_t \\ + \delta_{leader} \Delta P_{leader,t-1} + \delta_{other} \Delta P_{-i,t-1} + \alpha_X X_i + \alpha t + \varepsilon_{i,t},\end{aligned} \tag{I.3}$$

where $C_{i,t}$ is the dependent variable reflecting the own contribution in the current period t, L_t is the leader's contribution in the current period, $B_{i,t}$ is the belief imputed *from the conditional beliefs schedule from the one-shot game* that corresponds to the leader's contribution in period t, $H\text{-}leader$ and $L\text{-}follower$ are dummies reflecting the composition of returns in the group, $\Delta P_{leader,t-1}$ is the difference between the own payoff and the leader's payoff in the previous period, $\Delta P_{-i,t-1}$ is the difference between the own payoff and the payoff of the other follower in the group in the previous period, X_i is the set of individual characteristics, and $\varepsilon_{i,t}$ is a random effects error term.

For leaders' contributions, I estimate the following equation:

$$L_{j,t} = \omega_0 + \omega_1 H\text{-}leader + \omega_H \Delta P_{H,t-1} + \omega_L \Delta P_{L,t-1} + \omega t + \omega_X X_j + \eta_{j,t}, \tag{I.4}$$

where $L_{j,t}$ is the dependent variable reflecting the leader's contributions in the current period, t, $H\text{-}leader$ is a dummy reflecting the leader having a high return from the public good, $\Delta P_{H,t-1}$ is the difference between the own payoff and the payoff of H-follower in the previous period, $\Delta P_{L,t-1}$ is the difference between the own payoff and the payoff of L-follower in the previous period, X_j is the set of individual characteristics, and $\eta_{i,t}$ is a random effects error term. Equations (II.1) and (II.2) were estimated separately.

For the players in a repeated simultaneous game, I estimate (separately)

$$C_{k,t} = \beta_0 + \beta_1 \Delta P_{j\neq k,t-1} + \beta_2 \Delta P_{i\neq\{k,j\},t-1} + \beta t + \beta_L L\text{-}player + \beta_X X_k + \mu_{k,t}, \tag{I.5}$$

with $C_{k,t}$ being the own contribution in the current period, t, $\Delta P_{j\neq k,t-1}$ and $\Delta P_{k\neq\{j,k\},t-1}$ are the differences between players' own payoff and the payoff of the corresponding other players in the previous period. *L-player* is a dummy equal to 1 if the player has a low return from the public good, X_k is the set of individual characteristics, and $\mu_{k,t}$ is a random effects error term.

Table I.5 below shows the estimates for the dynamic panel random effects model for followers' contributions in a repeated sequential game. All standard errors are clustered at the group level (of three participants). The results in Table I.5 suggest that in their decision-making in period t followers who contribute after their leader in a repeated sequential game rely solely on the contribution of the leader in the current period. Differences in payoffs in the previous period are not statistically distinguishable from zero once additional individual controls are included. Overall, the contributions of followers decline over time and L-followers contribute on average about 3 points fewer compared with H-followers, although this difference is eliminated once additional controls are included. Similar to the one-shot game, followers seem to respond less generously to the contributions of H-leaders.

Result 8. *Followers' contributions in repeated sequential games are less sensitive to the contributions of H-leaders, similarly to their contribution preferences elicited in the one-shot game.*

Table I.6 shows estimates for a dynamic panel random effects model for leaders in a repeated sequential game. All standard errors are clustered at the group level (of three participants). The results in Table I.6 (column 2) suggest that the difference in payoffs between the leader and L-follower in a repeated sequential game is an important factor for leaders' contribution decisions in the current period. H-leaders contribute on average less than L-leaders, although leaders' contributions do not decline over time.

Result 9. *The average contributions of followers in a repeated sequential game decline over time. There is no such effect for the leaders.*

Table I.7 shows estimates for a dynamic panel random effects model for players in a repeated simultaneous game. All standard errors are clustered at the group level (of three participants). The results presented in Table I.7 show that differences in payoffs between players strongly matter for contribution decisions in the current period, but once individual controls are included only the difference in payoffs to one player remains important (which is, by the construction of variables, most often the L-player). All contributions decline over time.

Result 10. *Average contributions of players in a repeated simultaneous game decline over time.*

The results obtained in repeated games are in line with the findings in the literature on this topic. First, with repeated interaction, strong path dependence occurs. Leaders

Table I.5: Estimates for random effects model for dynamic contributions of followers (exogenous returns)

	(1) Contribution	(2) Contribution
Leader's contribution(t)	1.017***	1.045***
	(0.154)	(0.155)
Period, t	-0.460**	-0.484***
	(0.140)	(0.144)
Belief from one-shot game	-0.108	-0.148
	(0.100)	(0.130)
Difference with leader's profit($t-1$)	-0.122*	-0.080
	(0.061)	(0.094)
Difference with the profit of the other follower($t-1$)	-0.062	-0.054
	(0.054)	(0.040)
L-follower	-3.084*	-2.249
	(1.224)	(1.191)
H-leader	-1.040	0.064
	(2.176)	(2.125)
L-follower × Leader's contribution(t)	-0.064	-0.038
	(0.117)	(0.128)
H-leader × Leader's contribution(t)	-0.176	-0.196
	(0.146)	(0.121)
L-follower × H-leader	2.355	1.835
	(2.639)	(3.087)
L-follower × H-leader × Leader's contribution(t)	-0.190	-0.220
	(0.226)	(0.247)
Constant	6.094***	7.513***
	(1.838)	(1.930)
Other controls	No	Yes
Observations	270	270
Number of individuals	30	30

Notes: *L-follower* is a dummy equal to 1 if the follower has a low return from the public good; *H-leader* is a dummy equal to 1 if the leader in the group has a high return from the public good; $Difference\ with\ leader's\ profit(t-1)$ and $Difference\ with\ profit\ of\ the\ other\ follower(t-1)$ reflect differences in profits of the follower and respective players in the previous period. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

in repeated sequential games as well as players in repeated simultaneous games mostly rely on the differences in payoffs from the previous period. This effect is not observed for the followers in a repeated sequential game, who mainly rely on the composition of returns and leader's contribution. Initial beliefs do not seem to play a significant role in contribution decisions, at least in later periods (cf. Gächter and Renner, 2018). Second, overall contributions decline over time, which occurs, according to Fischbacher and Gächter (2010), due to the fact that players are imperfect conditional cooperators. Interestingly, I do not find the decline in leaders' contributions, which may suggest that leaders try to act as role models during the whole course of the experiment. Third, heterogeneity in

Table I.6: Estimates for random effects model for dynamic contributions of leaders (exogenous returns)

	(1) Contribution	(2) Contribution
Difference with profit of L-follower$(t-1)$	0.394***	0.402***
	(0.112)	(0.110)
Difference with profit of H-follower$(t-1)$	-0.211*	-0.264*
	(0.099)	(0.108)
H-leader	-4.426	-2.900
	(2.870)	(2.395)
Period, t	-0.238	-0.256
	(0.222)	(0.224)
Constant	14.420***	22.040***
	(2.164)	(2.898)
Other controls	No	Yes
Observations	135	135
Number of individuals	15	15

Notes: *H-leader* is a dummy equal to 1 if the leader in the group has a high return from the public good; $Difference\ with\ profit\ of\ L\text{-}follower(t-1)$ and $Difference\ with\ profit\ of\ H\text{-}follower(t-1)$ reflect differences in profits of the leader and respective followers in the previous period. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

Table I.7: Estimates for random effects model for dynamic contributions in simultaneous game (exogenous returns)

	(1) Contribution	(2) Contribution
Difference with the profit of player $1(t-1)$	-0.264*	-0.211
	(0.108)	(0.113)
Difference with the profit of player $2(t-1)$	0.285***	0.235*
	(0.081)	(0.093)
Period, t	-0.673*	-0.710*
	(0.284)	(0.294)
Lplayer	-1.019	-0.149
	(3.033)	(3.370)
Constant	14.308***	6.084
	(2.549)	(7.272)
Other controls	No	Yes
Observations	144	144
Number of individuals	16	16

Notes: *L-player* is a dummy equal to 1 if the player has a low return from the public good; $Difference\ with\ profit\ of\ player\ 1(t-1)$ and $Difference\ with\ profit\ of\ player\ 2(t-1)$ reflect differences in profits of the player and respective other players in the previous period. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

returns confirms affecting contributions in a repeated sequential game in a similar way as in the one-shot game, although, ceteris paribus, followers seem to be more responsive to leaders' contributions in a repeated interaction. Such responsiveness may occur due to either more precise belief updating in repeated games or better coordination on a particular contribution norm; however, I am unable to distinguish between these two hypotheses. Regarding the open empirical questions that arose from theoretical predictions, my findings suggest that at least in repeated games players in L-HL groups are able to achieve efficiency, giving up on inequality. According to the contribution dynamics patterns, five groups led by L-leaders were able to sustain a high level of cooperation until the last round of the repeated game, compared with only one group led by the H-leader. Moreover, the hypothesis on whether H-leaders and H-followers profit from cooperation regardless of the L-follower's contribution was not confirmed with my data, not even in repeated games.

I.7 Concluding Discussion

In this paper, I analyzed the effect of leading-by-example on the contributions in groups of three players who benefit differently from the public good. Leading-by-example is an institution that transforms a standard public goods game into a sequential one, where one player contributes first (the leader) and other players (the followers) contribute upon observing the leader's contribution. Heterogeneity in players' returns can give rise to normative conflicts due to the co-existence of multiple plausible rules, and leading-by-example seems to be a coordination device that potentially helps with the selection of a certain contribution rule. The experimental design allowed me to analyze the effect of heterogeneous returns on followers' conditional contributions, taking into account players' expectations about the contribution of the other follower in the group and individual characteristics such as gender, trust and risk preferences, and power distance. Additional treatments with repeated games also enabled me to study the effectiveness of leading-by-example compared with a standard voluntary contribution mechanism.

My main findings are as follows. The results obtained in the one-shot game indicate that contribution preferences change according to the composition of returns in the group. In particular, it is important what returns from the public good followers and leaders have: all followers seem to be less responsive to the contributions of H-leaders, but L-followers only respond to L-leaders, ceteris paribus. I establish that players demonstrate preferences for contribution rules that are similar to those reported in Reuben and Riedl (2013), and I further show that these preferences lead to differences in the effectiveness of an institution like leading-by-example. These differences most closely line up with the idea of a leader's contribution being perceived as a signal of the contribution norm: high contributions by L-leaders clearly exclude the possibility that the leader suggests any other contribution rule than contributing equally.

My results suggest that leading-by-example plays an important role in one-shot interactions. I find a strong impact of leaders' contributions on the contribution decisions of followers, which underlines the idea that leaders act as 'role models' (see Gächter and Renner, 2018; Glöckner et al., 2011, for earlier documentation of this effect), perhaps via shaping the beliefs of the followers in repeated interactions, which has a long-lasting effect on cooperation behavior. Nevertheless, I did not find statistical differences in contribution dynamics between repeated sequential and repeated simultaneous games, which suggests that strategic cooperation outweighs other effects like the effect of leading-by-example in my setup (see also Yamakawa et al., 2016). However, my findings provide evidence that players can sustain cooperation in groups with heterogeneous returns, overcoming inequality considerations.

Although my experimental design allowed me to avoid any reputation or strategic concerns, there remain some potential shortcomings that can confound the observed differences. I compared contributions in groups that differ only in the leader's returns from the public good. Future research is needed to understand whether the established differences are the result of the composition of returns in general, or if it is a property of following the example of the leader with a particular return: in my experiment, I keep the structure of followers' returns fixed, which is a rather restricting assumption. Furthermore, the results obtained in a laboratory experiment can be difficult to generalize. Further exploration is needed to establish whether the behavioral patterns discovered in this laboratory experiment are also observed in the field.

My results are complementary to the findings in Fischbacher et al. (2014): in their setup, they find no differences between conditional contributions of players with heterogeneous MPCRs and players with homogeneous MPCRs, while I establish heterogeneity in conditional contributions with respect to the composition of returns. Whether this heterogeneity leads to reduced cooperation compared with symmetric groups depends on the institutional environment (see, for example, Nikiforakis et al., 2012; Reuben and Riedl, 2013), whereby further investigation is needed to understand whether some environments may be less prone to negative effects caused by a missing common contribution norm than others.

The findings presented in this paper may be particularly interesting for behavioral theories of other-regarding preferences. While the contribution behavior in homogeneous groups is usually well explained by models of inequality aversion, the observed differences in followers' reactions to contributions of leaders with high or low returns cannot be fully explained by these models, as they do not capture the trade-off between equity and equality.

My results also suggest that changing the composition of returns among players may have indirect negative effects. For example, in societies where most members have low benefits from the public good, subsidies via increased benefits for some of the members may be confounded with the reduced likelihood of cooperation.

Heterogeneous reactions to the composition of returns underline the importance of belief management in groups with heterogeneous returns from the public goods (see also the discussion in Fischbacher et al., 2014). One can expect the signal obtained from low-benefiting leaders to be more valued by other group members than the one obtained from highly-benefiting leaders. For example, a researcher with high opportunity costs initializing a joint project with a high amount of effort most likely signals the high value or importance of this project. Contributions of people benefiting to a low extent from public facilities like public schools or public hospitals may induce higher involvement by bringing the importance of these facilities to the minds of people.

APPENDICES

I.A Experimental Instructions

These instructions were translated from German.

I.A.1 General Instructions for the Participants (paper)

You are now taking part in an economic study. If you read the explanations below carefully then you can — depending on your decisions and those of other participants — earn a not inconsiderable amount of money. It is therefore very important that you carefully read and understand these explanations.

During the study, there is an absolute ban on communication. If you have any questions, please address them to us only: raise your hand and we will come to you. Failure to comply with this rule will result in your exclusion from the trial and all payments.

The study comprises several independent parts. In the following, the course of the first part will be described first. Other parts will be described later. For today's participation, you will receive a show-up fee of 4 Euros. Your income is increased in each part of the study. However, during the study we are not talking about Euros, but about Talers. Your total income is thus initially calculated in Talers. The total number of Talers you earn during the study will then be converted into Euros and paid out in cash. The exchange rate from Talers to Euros will be announced at the beginning of the respective part.

I.A.2 Instructions for Part 1 (paper)

I.A.2.1 General Information

In the first part of the study, the exchange rate is: 10 Talers = 1 Euro.

The first part of the study comprises a **single period**. At the beginning of the first part, you are randomly divided into groups of three.

In H-HL groups:

In each group, there are two type B participants and one type A participant. If you are type B, your group will contain one other type B participant and one type A participant. If you are of type A, thus the other two participants are of type B.

In L-HL groups:

In each group, there are two type A participants and one type B participant. If you are type A, your group will contain one other type A participant and one type B participant. If you are of type B, thus the other two participants are of type A.

The types are randomly assigned and your type will be displayed soon on the monitor at the beginning of the first part.

Each participant receives 20 Talers. Your task now is to decide on the use of your 20 Talers. You can contribute all or part of the 20 Talers to a **project** or put it on a **private account**. Any Talers that you do not contribute to the project will be automatically added to your personal account. For example, if your contribution to the project is 5 Talers, 15 Talers will remain in your private account.

Income from the private account:
For every Taler that you put on the private account, you earn exactly one Taler. For example, if you place 20 Talers on the private account (and thus contribute nothing to the project), you will earn exactly 20 Talers from the private account. For example, if you contribute 12 Talers to the project (putting 8 Talers in your personal account), you will receive an income of 8 Talers from the private account. *No one except you receives any income from your private account.*

Income from the project:
For every Taler that you or another participant in your group contributes to the project, *you and all other participants in your group* receive a certain income from the project. How much income each participant receives from the project depends on which type he or she is.

Type A: income from the project = sum of contributions to the project × 0.5
Type B: income from the project = sum of contributions to the project × 0.75

In H-HL groups:
Examples: If the sum of the contributions of all participants to the project is 15 Talers (e.g. if you and the other two participants each contribute 5 Talers), then the participants of type B in your group will each receive $15 \times 0.75 = 11.25$ Talers from the project, and the participant of type A in your group receives $15 \times 0.5 = 7.5$ Talers. If instead the sum of the contributions to the project is 10 Talers in total, the participants of type B in your group will each receive $10 \times 0.75 = 7.5$ Talers from the project, and the participant of type A in your group will receive $10 \times 0.5 = 5$ Talers, etc.

In L-HL groups:
Examples: If the sum of the contributions of all participants to the project is 15 Talers (e.g. if you and the other two participants each contribute 5 Talers), then the participants of type A in your group will each receive $15 \times 0.5 = 7.5$ Talers from the project, and the participant of type B in your group receives $15 \times 0.75 = 11.25$ Talers. If instead the sum of the contributions to the project is 10 Talers in total, the participants of type A in your group will each receive $10 \times 0.5 = 5$ Talers from the project, and the participant of type B in your group will receive $10 \times 0.75 = 7.5$ Talers, etc.

Your pay-out from the first part is calculated by adding up your income from your personal account and your income from the project:

Type A:
Income from the private account (= 20 – contribution to the project)
+ Income from the project (= 0.5 × sum of contributions to the project)
Payment from the first part of the study

Type B:
Income from the private account (= 20 – contribution to the project)
+ Income from the project (= 0.75 × sum of contributions to the project)
Payment from the first part of the study

Procedure:
The contribution decisions are not made by all participants at the same time.

In H-HL groups:
Rather, one of the two participants of type B is randomly selected.

In L-HL groups:
Rather, one of the two participants of type A is randomly selected.

This participant is the first to make his contribution decision. Thereafter, the two other participants in the group will learn this contribution decision and then decide on their respective contribution decision.

Please answer the comprehension questions on the monitor. These should make it easier for you to familiarize yourself with the decision situation. The following table will help you to answer comprehension questions.

	Income from the project	
Sum of contributions to the project	Type A	Type B
20	10	15
40	20	30
60	30	45

If you still have questions, please hold a hand out of the cabin - we will come to you.

I.A.2.2 Additional Explanations for the First Part of the Study

Your contribution decision will be made in the first part of the upcoming study.

In H-HL groups:
In each group, a randomly-selected type B participant makes a contribution decision first.

In L-HL groups:
In each group, a randomly-selected type A participant makes a contribution decision first.

Before the decision of the first participant is communicated to the other two participants, they must complete a contribution table. In the contribution table, you have to decide for each of the 21 possible contribution decisions of the first participant how many

Talers you want to contribute to the project yourself. For a better understanding, look at the screenshot in Figure I.A.1 below, which you will see later on the monitor in case you are *not* the first to decide on contribution:

Ihre Anfangsausstattung: 20, Ihr Ertrag (je Taler) aus dem Projekt: 0.50

Ertrag (je Taler) des Teilnehmers, der zuerst seine Beitragsentscheidung gewählt hat: 0.75

Ertrag (je Taler) des anderen Teilnehmers, der zeitgleich mit Ihnen seine Beitragsentscheidung wählt: 0.75

In der folgenden Tabelle stehen links die möglichen Beitragsentscheidungen des Teilnehmers, welcher als erster seine Beitragsentscheidung wählt (von 0 bis 20).
In das Feld daneben tragen Sie ein, wie hoch **Ihre eigene Beitragsentscheidung** in diesem Fall sein soll.
Drücken Sie auf "OK", wenn Sie fertig sind.

Die anderen Teilnehmer werden erst am Ende der gesamten Studie erfahren, wie Sie sich entscheiden haben.

Beitragsentscheidungen des ersten Teilnehmers	Ihr Beitrag zum Projekt	Beitragsentscheidungen des ersten Teilnehmers	Ihr Beitrag zum Projekt	Beitragsentscheidungen des ersten Teilnehmers	Ihr Beitrag zum Projekt
0		7		14	
1		8		15	
2		9		16	
3		10		17	
4		11		18	
5		12		19	
6		13		20	

Figure I.A.1: Screenshot with contribution table in German (exogenous returns)

The numbers (on the screenshot) to the left of the blue input fields are the possible contributions of the participant who first defined his or her contribution decision. You now simply have to enter in each input field how many Talers *you* want to contribute to the project, provided that the first participant provides the specified contribution. **You have to make an entry in each field.**

For example, you must specify how much you contribute to the project when the first participant contributes 0 Talers to the project; how much you contribute when the first participant contributes 1 Taler, or 2 or 3, and so on. You can enter integers from 0 to 20 in each field for **all** fields. Since you do not know at the time of the 21 decisions which of them will be relevant, you should take each one of the decisions carefully.

Only at the end of the study, you will learn what the actual contribution decision of the first participant was. This decision and the decisions of the other two participants for this situation will eventually determine the pay-outs that participants will receive at the end of the first part.

Do you have more questions? If so, please hold a hand out of the cabin.

I.A.3 Instructions for Part 2 (from the Computer Screen)

I.A.3.1 General Information About the Second Part of the Study

The second part of the study comprises a total of *ten periods*. The decision situation in each period is similar to the previous situation: you are back in the same trio as before.

In H-HL groups:

There are each again two type B participants and one type A participant.

In L-HL groups:

There are each again two type A participants and one type B participant.

They are of the same type as before.

Each participant receives 20 Talers at the beginning of *each* period. Your job is again, to decide how many Talers you want to contribute to a project (the rest remains back in your private account).

In treatments SEQ:

The contribution decisions are not made by all participants at the same time.

In H-HL groups:

Again, one of the two participants of type B (the same participant as in the part before) will be the first to make a contribution decision.

In L-HL groups:

Again, one of the two participants of type A (the same participant as in the part before) will be the first to make a contribution decision.

After that, the other two participants in the group will learn this contribution decision and then make their own contribution decision. **Unlike before**, these two participants do not have to enter their decision in the form of a contribution table. Instead, they learn directly the actual decision of the first participant and therefore also enter only one contribution decision, which is immediately pay-out relevant.

In treatments SIM:

Unlike before, the contribution decisions are now made by **all** participants at the **same time**. **Unlike before**, the decision does not have to be entered in the form of a contribution table. Instead, all participants enter only one contribution decision, which immediately becomes pay-out relevant.

At the end of each period, you will learn about the contributions and pay-outs of all participants in your group. Then the next period begins. Your type and composition of the group remains the same over all ten periods. Likewise, the same participant is the first to enter his or her contribution decision.

The amount of pay-outs per period is determined as before. In the end, you will receive the sum of your pay-outs over all ten periods, with the exchange rate from Euro to Talers now: 1 Taler = 0.02 Euro.

If you have any questions, please hold a hand out of the cabin, and we will come to you.

I.B Theoretical Predictions with Social Preferences

Without loss of generality I assume that MPCRs are such that $\gamma_H = (1+\delta)\gamma_L \equiv (1+\delta)\gamma$, $\delta > 0$. In what follows I will use subscripts i and j to refer to utility and parameters of the followers, and subscript l to refer to variables related to the leader. In cases when it is necessary to distinguish between players with high and low returns, H- and L-players, I will use subscripts H and L.

I.B.1 Fehr-Schmidt Preferences

I analyze the equilibria in the sequential linear public goods game with three players under the assumptions on the shape of players' utility function: each player's utility function coincides with the monetary payoff π_i from the public goods experiment, and players have social preferences, namely they suffer from disutility from the inequality of monetary payoffs, as defined in Fehr and Schmidt (1999).

Although there are known drawbacks of linear Fehr-Schmidt Model (the marginal rate of substitution between monetary income and inequality is constant), it is simple to solve and is useful to demonstrate the key differences in players' behavior.

The preferences of players are given by

$$u_i(x_1,\ldots,x_n) = \pi_i(\vec{x}) - \frac{\alpha_i}{n-1}\sum_{j\neq i}\max\{0,\pi_j(\vec{x})-\pi_i(\vec{x})\} - \frac{\beta_i}{n-1}\sum_{j\neq i}\max\{0,\pi_i(\vec{x})-\pi_j(\vec{x})\}, \tag{I.1}$$

where $\alpha_i \geq \beta_i$ and $0 < \beta_i \leq 1$ are the parameters of the utility function, and π_i is the monetary payoff from the public goods experiment. Variables x_i represent individual contribution decisions in the public goods game, and $\vec{x} = (x_1, x_2, \ldots, x_n)$ denotes the vector of the decision of all n players in the group.

Suppose every player has FS preferences (I.1). I start with H-HL groups and I first analyze equilibria in a simultaneous game. I analyze equilibria in a sequential game second. I find equilibria in the subgame in the second stage of the sequential game, and conditions under which these equilibria may occur. After that, I analyze the behavior of the leader given the equilibrium conditions for the second stage. Then I continue with L-HL groups.

I.B.1.1 H-HL Group

Simultaneous Game

The predictions for the simultaneous game in H-HL group are absolutely identical to the predictions obtained in Kube et al. (2015) for HET – VCM treatment (see Kube et al., 2015, Proposition 2 in the appendix).

Proposition 1. *In simultaneous game in the group with two players with high return and* *one player with low return, if for both H-players the following inequality is satisfied*

$$\beta_H > 2\frac{1-\gamma(1+\delta)}{2-\delta\gamma}, \tag{I.2}$$

and for the L-player

$$\beta_L > \frac{1-\gamma}{1+\delta\gamma} \tag{I.3}$$

is satisfied, then there exist equilibria in which all players contribute positive amounts: H-players contribute $c_H \in (0;20]$ *and L-player contributes* $c_L = c_H\frac{1-2\delta\gamma}{1+\delta\gamma}$. All players receive equal payoffs.

If for both H-players the following inequality is satisfied

$$2\gamma(1+\delta) + \beta_i \geq 2 + \alpha_i(1-\delta\gamma), \tag{I.4}$$

but for L-player (I.3) is not satisfied, then there exist equilibria in which L-player contributes zero, but H-players contribute equal positive amounts $c_H \in (0;20]$. Players receive unequal payoffs.[36]

Otherwise, in the unique Nash equilibrium all players contribute zero to the public good.

For the parameters in my experiment, conditions (I.2) and (I.3) are simplified into $\beta_H > \frac{2}{7}$ and $\beta_L > \frac{2}{5}$, *only equilibria with equal payoffs exist*, and in equilibria with positive contributions H-players contribute $c_H \in (0;20]$ and L-player contributes $c_L = 2c_H/5$.

Sequential Game

I start with the L-follower. Suppose that the leader contributes L and the H-follower contributes F. Then the utility function of L-follower is then given by

$$\begin{aligned} u_i(x,L,F) = 20 - x + \gamma(x+L+F) \\ &- \frac{\alpha_i}{2}\big[\max\{0, x - F + \delta\gamma(x+L+F)\} + \max\{0, x - L + \delta\gamma(x+L+F)\}\big] \\ &- \frac{\beta_i}{2}\big[\max\{0, F - x - \delta\gamma(x+L+F)\} + \max\{0, L - x - \delta\gamma(x+L+F)\}\big]. \end{aligned}$$

There are two thresholds that are relevant for utility maximization

$$T_l = \frac{L - \delta\gamma(F+L)}{1+\delta\gamma}, \quad \text{and} \quad T_f = \frac{F - \delta\gamma(F+L)}{1+\delta\gamma}.$$

These thresholds represent the contribution levels, required to make the payoff of the L-follower equal the payoff of the leader and the payoff of another follower, given that the leader contributes L and another follower contributes F.

[36] These equilibria only exist for very specific returns from the public good. For example, the extreme case $\alpha_i = \beta_i$ transforms condition (I.4) into $\beta_i \geq \frac{2-2\gamma(1+\delta)}{\delta\gamma}$ which, together with $\beta_i \leq 1$ and $\gamma(1+\delta) < 1$, give restrictions on the returns.

The L-follower contributes positive amount if condition (I.3) is satisfied, and never contributes more than required to make her payoff equal the largest payoff among other players: either the leader's payoff (and, thus, $c_L = T_l$), or the payoff of another follower (and, thus, $c_L = T_f$), due to the restriction on inequality aversion parameters α and β, $\alpha \geq \beta$, and returns from the public good γ, $\gamma < 1$.

For the H-follower the utility is given by

$$\begin{aligned} u_i(x, L, F) = 20 - x + \gamma(1+\delta)(x+L+F) \\ - \frac{\alpha_i}{2}\left[\max\{0, x-L\} + \max\{0, x-F-\delta\gamma(x+L+F)\}\right] \\ - \frac{\beta_i}{2}\left[\max\{0, L-x\} + \max\{0, F-x+\delta\gamma(x+L+F)\}\right]. \end{aligned}$$

For the H-follower the contribution level required to make her payoff equal to the payoff of the leader is simply L, and contribution level required to make her payoff equal the payoff of another follower is

$$T = \frac{F(1+\delta\gamma) + \delta\gamma L}{1-\delta\gamma}.$$

The H-follower contributes positively if

$$\beta_H > \frac{1-\gamma(1+\delta)}{2-\delta\gamma}. \tag{I.5}$$

Note that the H-follower can theoretically contribute more than required to make her payoff equal to the payoff of another follower (L-follower), in order to match the payoff of the leader (thus, $c_H = L$). This happens if condition (I.4) is satisfied.

Suppose that the leader is fully informed, i.e. he knows the inequality aversion of the followers (their α_i and β_i). Then there are two possible cases:

Case 1: both followers contribute positively if and only if the leader contributes positively (inequalities (I.3) and (I.5) are satisfied, but (I.4) is not);

Case 2: the H-follower contributes the same amount the leader contributes, regardless of the contribution of another follower (inequality (I.4) is satisfied).

Case 1: Suppose the followers are such that they contribute positive amounts only if leader contributes, and they contribute $c_H = F \leq L$ and $c_L = T_f = \frac{F(1-\delta\gamma)-\delta\gamma L}{1+\delta\gamma}$, correspondingly. In this case, the leader suffers from disadvantageous inequality

$$\begin{aligned} u_l(L, F) &= 20 - L + \gamma(1+\delta)\left(L + F + \frac{F(1-\delta\gamma) - \delta\gamma L}{1+\delta\gamma}\right) - \alpha_l(L-F) \\ &= 20 - L + \gamma(1+\delta)\frac{L+2F}{1+\delta\gamma} - \alpha_l(L-F) \end{aligned}$$

On the one hand, the leader should contribute $L > 0$ if

$$\gamma(1+\delta)\frac{L+2F}{1+\delta\gamma} \geq L + \alpha_l(L-F),$$

for given F. On the other hand, the leader is strictly better off if she reduces the differences between her payoff and the payoff of H-follower, thus, $L = F$ as soon as $F > 0$, and all players receive equal payoffs.

Case 2: If the H-follower always contributes the same amount the leader does, then the leader can still suffer from disadvantageous inequality if the L-follower contributes $0 \leq F < L\frac{1-2\delta\gamma}{1+\delta\gamma}$. However, the leader always contributes the full endowment if

$$\alpha_l \leq \frac{4\gamma(1+\delta)-2}{1-2\delta\gamma}. \tag{I.6}$$

Proposition 2. *In a sequential public goods game in H-HL group:*

- *If either* (I.3) *for the L-follower or* (I.5) *for the H-follower are* not *satisfied, then in the unique subgame perfect equilibrium all players contribute zero to the public good.*
- *If both* (I.3) *for the L-follower, and* (I.5) *for the H-follower are satisfied, then there are equilibria, in which the leader contributes L^* (up to the full endowment), the H-follower contributes $c_H = L^*$ and the L-follower contributes $c_L = \frac{(1-2\delta\gamma)L^*}{1+\delta\gamma}$.*
- *If* (I.4) *for the H-follower and* (I.6) *for the leader are satisfied, but* (I.3) *for the L-follower is* not *satisfied, then there are equilibria in which the leader and the H-follower contribute the full endowment, and the L-follower contributes zero.*

For the parameters in my experiment, conditions (I.3) and (I.5) are simplified into $\beta_L > \frac{2}{5}$ and $\beta_H > \frac{1}{7}$, *only equilibria with equal payoffs exist*, and in all equilibria the leader contributes L, the H-follower contributes $c_H = L$, and the L-follower contributes $c_L = 2L/5$.

I.B.1.2 L-HL Group

Simultaneous Game

Proposition 3. *In a simultaneous game in the group with two L-players and one H-player, if for both L-players the following inequality holds*

$$\beta_L > 2\frac{1-\gamma}{2+\delta\gamma}, \tag{I.7}$$

and if for the H-player it holds that

$$\beta_H > \frac{1-\gamma(1+\delta)}{1-\delta\gamma}, \tag{I.8}$$

then there exist equilibria in which all players contribute positive amounts: the H-player contributes $c_H \in (0;20]$ *and L-players contribute* $c_L = c_H \frac{1-\delta\gamma}{1+2\delta\gamma}$*,* all players receive equal payoffs*; otherwise, in the unique Nash equilibrium all players contribute zero to the public good.*

For the parameters in my experiment, conditions (I.7) and (I.8) are simplified into $\beta_L > \frac{4}{9}$ and $\beta_H > \frac{1}{3}$, and in equilibria with positive contributions the H-player contributes $c_H \in (0;20]$ and L-players contribute $c_L = c_H/2$.

Sequential Game

I start with the L-follower. Suppose, the leader contributes L and another follower (H-follower) contributes F. Then the utility function of L-follower is given by

$$\begin{aligned} u_i(x,L,F) = 20 - x + \gamma\,[x+L+F] \\ - \frac{\alpha_i}{2}\,[\max\{0, x-L\} + \max\{0, x-F+\delta\gamma(x+L+F)\}] \\ - \frac{\beta_i}{2}\,[\max\{0, L-x\} + \max\{0, F-x-\delta\gamma(x+L+F)\}]\,. \end{aligned}$$

For the L-follower contribution level required to make her payoff equal to the payoff of the leader is L, and contribution level required to make her payoff equal to the payoff of another follower is

$$T' = \frac{(1-\delta\gamma)F - \delta\gamma L}{1+\delta\gamma}.$$

The L-follower contributes positively if

$$\beta_L > \frac{1-\gamma}{2+\delta\gamma}. \tag{I.9}$$

If this condition is satisfied, she will contribute the minimal amount required to make her payoff equal to either the leader's payoff (and, thus, $c_L = L$), or the payoff of another follower (and, thus, $c_L = T'$), whichever is higher.

Now, consider the H-follower. Suppose, the leader contributes L and another follower (L-follower) contributes F. Then the utility function of the H-follower is given by

$$\begin{aligned} u_i(x,L,F) = 20 - x + \gamma(1+\delta)(x+L+F) \\ - \frac{\alpha_i}{2}\big[\max\{0, x-F-\delta\gamma(x+L+F)\} + \max\{0, x-L-\delta\gamma(x+L+F)\}\big] \\ - \frac{\beta_i}{2}\big[\max\{0, F-x+\delta\gamma(x+L+F)\} + \max\{0, L-x+\delta\gamma(x+L+F)\}\big]. \end{aligned}$$

For the H-follower there are now two thresholds

$$T'_l = \frac{L+\delta\gamma(F+L)}{1-\delta\gamma}, \quad \text{and} \quad T'_f = \frac{F+\delta\gamma(F+L)}{1-\delta\gamma},$$

which correspond with the contributions required to equalize payoffs with the leader and with another follower, correspondingly. Note that both thresholds are increasing in both L and F.

The H-follower contributes a positive amount if condition (I.8) is satisfied. If this condition is satisfied, she will contribute the minimal amount required to make her payoff equal to either the leader's payoff (and, thus, $c_H = T'_l$), or the payoff of another follower (and, thus, $c_H = T'_f$), whichever is higher.

Suppose that the leader is fully informed, i.e. she knows the inequality aversion of the followers (their β_i). Suppose that the followers are such that they contribute in equilibrium $c_L = F \leq L$ and $c_H = T'_f = \frac{F(1+\delta\gamma)F+\delta\gamma L}{1-\delta\gamma}$, correspondingly. Then the leader suffers from disadvantageous inequality

$$\begin{aligned} u_l(L, F(L)) = 20 - L + \gamma\left[L + F + c_H(F)\right] \\ - \alpha_l\left[(L - F) + (L - c_H(F) + \delta\gamma(F + c_H(F) + L))\right]/2 \\ = 20 - L + \gamma\left[L\frac{1}{1-\delta\gamma} + F\frac{2}{1-\delta\gamma}\right] - \alpha_l(L - F). \end{aligned}$$

On the one hand, the leader should contribute $L > 0$ if

$$\gamma\left[L + 2F\right] \geq (1 - \delta\gamma)[L + \alpha_l(L - F)],$$

for given F. On the other hand, the leader is strictly better off if she reduces the differences between her payoff and the payoff of L-follower, thus, $L = F$ as soon as $F > 0$, and all players receive equal payoffs.

If the leader contributes more than the amount, required to equalize payoffs when the H-follower contributes $c_H = 20$ and the L-follower contributes $c_L = \frac{20(1-\delta\gamma)-\delta\gamma L}{1+\delta\gamma}$, she will have the lowest payoff in the end. Since the leader knows that the followers choose contributions that equalize their payoffs with the highest payoff in the group and that the H-follower has bounded endowment, therefore they never contribute more than $c_H = 20$ and $c_L = \frac{20(1-\delta\gamma)-\delta\gamma L}{1+\delta\gamma}$. Thus, the leader never contributes more than $\bar{E} = 20\frac{1-\delta\gamma}{1+2\delta\gamma}$.

Proposition 4. *In a sequential public goods game in L-HL group:*

- *If either* (I.8) *for the H-follower or* (I.9) *for the L-follower are* not *satisfied, then in the unique subgame perfect equilibrium all players contribute nothing.*

- *If both* (I.8) *for the H-follower and* (I.9) *for the L-follower are satisfied, then there are equilibria where the leader contributes L^* (up to $\bar{E} = 20\frac{1-\delta\gamma}{1+2\delta\gamma}$), and the followers contribute $c_L = L^*$, and $c_H = \frac{(1+2\delta\gamma)L^*}{1-\delta\gamma}$.*

For the parameters in my experiment, conditions (I.8) and (I.9) are simplified into $\beta_H > \frac{1}{3}$ and $\beta_L > \frac{2}{9}$ and in equilibria with positive contributions the leader contributes L, the L-follower contributes $c_L = L$, and the H-follower contributes $c_H = 2L$.

I.B.2 Charness and Rabin Preferences

I analyze the equilibria in the sequential linear public goods game with three players under the assumption that players have social welfare preferences, as defined in the Appendix 1 of Charness and Rabin (2002) (i.e. absent of reciprocity).

The preferences of player i are given by

$$U_i(\pi_1, \ldots, \pi_n) = (1-\lambda)\pi_i + \lambda \left[\sigma \min\{\pi_1, \ldots, \pi_n\} + (1-\sigma) \sum_{j=1}^{n} \pi_j \right], \tag{I.10}$$

where π_i is player i's monetary payoff from the public goods experiment, $\lambda \in [0,1]$ measures how much player i cares about pursuing the social welfare versus her own self-interest, and $\sigma \in [0,1]$ is a parameter measuring the degree of concern for helping the worst-off person versus maximizing the total social surplus.

Suppose every player has CR preferences (I.10). I start with H-HL groups and I first analyze equilibria in a simultaneous game. I analyze equilibria in a sequential game second. I find equilibria in the subgame in the second stage of the sequential game, and conditions under which these equilibria may occur. After that, I analyze the behavior of the leader given the equilibrium conditions for the second stage. Then I continue with L-HL groups.

I.B.2.1 H-HL Group

Simultaneous Game

A priori there is no reason for an L-player to believe H-players to contribute unequally. Thus, the utility function of the L-player is transformed into

$$\begin{aligned} U_L(x, x_H) = {} & (1-\lambda_L)\left[20 - x + \gamma(2x_H + x)\right] \\ & + \lambda_L \left[\sigma_L \min\left\{20 - x_H + \gamma(1+\delta)(2x_H + x), 20 - x + \gamma(2x_H + x)\right\}\right. \\ & \left. + (1-\sigma_L)\left(60 + [\gamma(3+2\delta) - 1](2x_H + x)\right)\right]. \end{aligned}$$

Marginal utility here is a step function and its sign depends on the parameters of the utility function λ and σ and on the contribution of the H-player:

$$\frac{\partial u_L}{\partial x} = \begin{cases} \gamma - 1 + \lambda_L \gamma[2 + 2\delta - \sigma_L(2+2\delta)], & \text{if } x > \frac{1-2\delta\gamma}{1+\delta\gamma} x_H, \\ \gamma - 1 + \lambda_L[\gamma(2+2\delta) - \sigma_L(\gamma(2+\delta) - 1)], & \text{otherwise.} \end{cases}$$

It is always positive if

$$\lambda_L > \frac{1-\gamma}{\gamma(2+2\delta)(1-\sigma_L)} \tag{I.11}$$

and

$$\lambda_L > \frac{1-\gamma}{\gamma(2+2\delta) - \sigma_L[\gamma(2+\delta) - 1]} \tag{I.12}$$

are satisfied, then the L-player always wants to contribute as much as possible (due to her taste of social welfare).[37]

Given the parameters chosen in my design ($\gamma = 0.5$ and $\delta = 0.5$), the condition takes form

$$\lambda_L > \frac{1}{3 - 3\sigma_L}.$$

For H-players, the utility function is transformed into (assuming that the L-player contributes x_L and another H-player contributes x_H)

$$\begin{aligned} U_H(x, x_H, x_L) = &(1-\lambda_H)\left[20 - x + \gamma(1+\delta)(x + x_H + x_L)\right] \\ &+ \lambda_H \left[\sigma_H \min\{20 - x + \gamma(1+\delta)(x + x_H + x_L), \right. \\ &\quad 20 - x_H + \gamma(1+\delta)(x + x_H + x_L), \\ &\quad 20 - x_L + \gamma(x + x_H + x_L)\} \\ &\left. + (1-\sigma_H)\left(60 + [\gamma(3+2\delta) - 1](x + x_H + x_L)\right)\right]. \end{aligned}$$

There is a threshold in H-players' contributions that makes payoffs of different players equal: if the contribution of an H-player equals T, then her payoff equals the payoff of an L-player (assume the contribution of another H-player is z):

$$T(z) = \frac{(1+\delta\gamma)x_L - \delta\gamma z}{1 - \delta\gamma}.$$

$$\frac{\partial u_H}{\partial x} = \begin{cases} \gamma(1+\delta) - 1 + \lambda_H[\gamma(2+\delta)(1-\sigma_H)], & \text{if } x > x_H,\ x > T(x_H), \\ \gamma(1+\delta) - 1 + \lambda_H[\gamma(2+\delta) - \sigma_H(\gamma(2+\delta) - 1)], & \text{if } x < x_H, T(x) < x_H, \\ \gamma(1+\delta) - 1 + \lambda_H[\gamma(2+\delta) - \sigma_H(\gamma(2+2\delta) - 1)], & \text{if } x < T(x_H),\ x_H < T(x). \end{cases}$$

This marginal utility is always positive if

$$\lambda_H > \frac{1 - \gamma(1+\delta)}{\gamma(2+\delta)(1-\sigma_H)}, \tag{I.13}$$

$$\lambda_H > \frac{1 - \gamma(1+\delta)}{\gamma(2+\delta) - \sigma_H(\gamma(2+\delta) - 1)} \tag{I.14}$$

and

$$\lambda_H > \frac{1 - \gamma(1+\delta)}{\gamma(2+\delta) - \sigma_H(\gamma(2+2\delta) - 1)} \tag{I.15}$$

are satisfied, then the H-player always wants to contribute as much as possible (due to her taste of social welfare). Again, the last two inequalities are redundant.

[37] It is clear that the second inequality is redundant: if the first one is satisfied, then the second one is also satisfied, given the assumptions on the parameters of utility function, $\lambda \in [0;1]$ and $\sigma \in [0;1]$. The second inequality is only important when the first one is violated, then if the second one is satisfied, it gives limitations on parameters' values that lead to different equilibrium.

Given the parameters chosen in my design, the condition takes form

$$\lambda_H > \frac{1}{5 - 5\sigma_H}.$$

However, if for an L-player condition (I.11) is satisfied, then she receives the lowest payoff by contributing the full endowment, regardless what other players contribute. Thus, it is sufficient for only condition (I.15) to be satisfied for L-players to contribute the full endowment.

Sequential Game

Suppose that in a sequential game the L-follower observes the leader's contribution L, and the H-follower contributes x_H. Her utility is then given by

$$\begin{aligned} U_L(x, L, x_H) = (1 - \lambda_L)\,&[20 - x + \gamma(x + L + x_H)] \\ &+ \lambda_L\,[\sigma_L \min\{20 - x + \gamma(x + L + x_H), \\ &\qquad 20 - L + \gamma(1+\delta)(x + L + x_H), \\ &\qquad 20 - x_H + \gamma(1+\delta)(x + L + x_H)\} \\ &\quad + (1 - \sigma_L)\,(60 + [\gamma(3 + 2\delta) - 1](x + L + x_H))]\,. \end{aligned}$$

Marginal utility is similar to the case of simultaneous game,

$$\frac{\partial u_L}{\partial x} = \begin{cases} \gamma - 1 + \lambda_L[\gamma(2+2\delta)(1-\sigma_L)], & \text{if } x > \frac{(1-\delta\gamma)L - \delta\gamma x_H}{1+\delta\gamma}, \\ (1-\lambda_L)(\gamma - 1) + \lambda_L[\gamma(2+2\delta) - \sigma_L(\gamma(2+\delta) - 1)], & \text{otherwise,} \end{cases}$$

so it is always positive if conditions (I.11) is satisfied (again, condition (I.12) is redundant), so the L-follower is always willing to contribute the full endowment.

Similarly, the H-follower is always willing to contribute the full endowment if condition (I.13) is satisfied. At the same time, if the L-follower contributes the full endowment, she always receives the lowest payoff, so for the H-follower it is sufficient that condition (I.15) is satisfied to contribute the full endowment.

The same applies for the H-leader: she is always willing to contribute the full endowment if condition (I.13) is satisfied, but once the L-follower contributes the full endowment, she always receives the lowest payoff, so for the leader it is sufficient that condition (I.15) is satisfied to contribute the full endowment.

Proposition 5. *In both simultaneous and sequential games, in the group with two players with high return and one player with low return, if conditions (I.11) for an L-player and (I.15) for H-players are satisfied, then there exists unique equilibrium in which all players contribute the full endowment.*

If neither condition (I.12) for an L-player nor condition (I.15) for H-players are satisfied, then in the unique Nash equilibrium all players contribute zero to the public good.

I.B.2.2 L-HL Group

Simultaneous Game

A priori there is no reason for an H-player to believe L-players to contribute unequally. Thus, the utility function of the H-player is transformed into

$$\begin{aligned} U_H(x, x_L) = &(1-\lambda_H)\left[20 - x + \gamma(1+\delta)(2x_L + x)\right] \\ &+ \lambda_H\left[\sigma_H \min\left\{20 - x + \gamma(1+\delta)(2x_L + x), 20 - x_L + \gamma(2x_L + x)\right\}\right. \\ &\left. +(1-\sigma_H)\left(60 + [\gamma(3+\delta) - 1](2x_L + x)\right)\right]. \end{aligned}$$

Marginal utility is

$$\frac{\partial u_H}{\partial x} = \begin{cases} \gamma(1+\delta) - 1 + 2\gamma\lambda_H(1-\sigma_H), & \text{if } x > \frac{1+2\delta\gamma}{1-\delta\gamma} x_L, \\ \gamma(1+\delta) - 1 + \lambda_H[2\gamma - \sigma_H\gamma(2+\delta) + \sigma_H], & \text{otherwise.} \end{cases}$$

and is always positive (the H-player is always willing to contribute the full endowment) if

$$\lambda_H > \frac{1-\gamma(1+\delta)}{2\gamma(1-\sigma_H)} \tag{I.16}$$

and

$$\lambda_H > \frac{1-\gamma(1+\delta)}{2\gamma - \sigma_H(\gamma(2+\delta)-1)} \tag{I.17}$$

are satisfied (the second inequality is redundant).

For L-players, the utility function takes the form (assuming that the H-player contributes x_H and another L-player contributes x_L)

$$\begin{aligned} U_L(x, x_H, x_L) = &(1-\lambda_L)\left[20 - x + \gamma(x + x_H + x_L)\right] \\ &+ \lambda_L\left[\sigma_L \min\left\{20 - x + \gamma(x + x_H + x_L),\right.\right. \\ &\quad 20 - x_H + \gamma(1+\delta)(x + x_H + x_L), \\ &\quad \left. 20 - x_L + \gamma(x + x_H + x_L)\right\} \\ &\left. +(1-\sigma_L)\left(60 + [\gamma(3+\delta) - 1](x + x_H + x_L)\right)\right]. \end{aligned}$$

Marginal utility depends on a threshold in the contribution of L-players

$$T(z) = \frac{(1-\delta\gamma)x_H - \delta\gamma z}{1+\delta\gamma},$$

that makes their payoff equal the payoff of the H-player. Therefore,

$$\frac{\partial u_L}{\partial x} = \begin{cases} \gamma - 1 + \lambda_L\gamma(2+\delta)(1-\sigma_L), & \text{if } x > x_L,\ x > T(x_L), \\ \gamma(1+\delta) - 1 + \lambda_L[2\gamma - (2\gamma-1)\sigma_L], & \text{if } x < T(x_L),\ x_L < T(x), \\ \gamma - 1 + \lambda_L[\gamma(2+\delta) - (\gamma(2+\delta)-1)\sigma_L], & \text{if } x < x_L, T(x) < x_L. \end{cases}$$

It is always positive (the L-player is always willing to contribute the full endowment) if

$$\lambda_L > \frac{1-\gamma}{\gamma(2+\delta)(1-\sigma_L)}, \tag{I.18}$$

$$\lambda_L > \frac{1-\gamma}{2\gamma-(2\gamma-1)\sigma_L} \tag{I.19}$$

and

$$\lambda_L > \frac{1-\gamma}{\gamma(2+\delta)-[\gamma(2+\delta)-1]\sigma_L} \tag{I.20}$$

are satisfied (again, the last two conditions are redundant).

If one of L-players contributes the full endowment (condition (I.18) satisfied), then she receives the lowest payoff, regardless what other players contribute (only if another L-player contributes the full endowment, they both receive the lowest payoffs). Thus, it is sufficient for only condition (I.17) to be satisfied for an H-player, and condition (I.20) to be satisfied for another L-player to make them contribute their full endowment.

Sequential Game

Equilibrium conditions in a sequential game are similar to those in a simultaneous game: if condition (I.17) is satisfied, then (given that one of L-players, either the leader or the L-follower, contribute the full endowment) the H-follower is also willing to contribute the full endowment, since her payoff is greater than of other players. If one of L-players, either the leader or the L-follower, have parameters of utility function that satisfy condition (I.18), then she is always willing to contribute the full endowment, regardless what others contribute. Then, if for the L-follower condition (I.20) is satisfied, she is also willing to contribute the full endowment, since her payoff then is not lower that of another L-player, the leader.

Proposition 6. *In both simultaneous and sequential games, in the group with two players with low return and one player with high return, if conditions* (I.17) *for an H-player,* (I.18) *for one L-player (either the leader or the L-follower in a sequential game), and* (I.20) *for another L-player (the L-follower / the leader in a sequential game) are satisfied, then there exists unique equilibrium in which all players contribute the full endowment.*

If none of condition (I.17) *for an H-player, and condition* (I.20) *for L-players are satisfied, then in the unique Nash equilibrium all players contribute zero to the public good.*

One should note that there might also exist equilibria in which H-players contribute the full endowment, and L-players contribute such amounts that make their payoffs equal to payoffs H-followers, e.g. $\frac{1-2\delta\gamma}{1+\delta\gamma}20$ in H-HL groups and $\frac{1-\delta\gamma}{1+2\delta\gamma}20$ in L-HL groups (i.e. 8 and 10 for the parameters chosen in my design).

I.B.3 Adherence to Fairness Ideals

I analyze the equilibria in the sequential linear public goods game with three players under the assumptions on the shape of players' utility function: each player's utility function depends on the adherence to a fairness ideal/social norm. More formally, a player's utility function coincides with monetary payoff, and every player suffers from disutility from non-compliance with fairness ideals/social norms (by experiencing quadratic loss from deviations).[38]

The utility functions of players are given by

$$V_i(c_i, c_{-i}) = \pi(c_i, c_{-i}) - \frac{\beta_i}{2}\left(c_i - m^{k(i)}(c_{-i})\right)^2, \tag{I.21}$$

where $c_{-i} = \sum_{j \neq i}^{n} c_j$, the parameter $\beta \geq 0$ determines the weight attached to deviations from the fairness ideal $m^{k(i)}(c_{-i})$, and index $k(i)$ represents the ideal preferred by player i, $k \in \{eq, equi\}$.

I distinguish between the following fairness ideals: equality suggests that players value the equalization of absolute contributions of other players with no links to individual characteristics such as MPCR from the public good:

$$m^{eq} = \frac{1}{n-1} c_{-i}.$$

Equity suggests that players value the equalization of payoffs of other players. Given the payoff structure of the public goods game, this leads to

$$m^{equi} = \frac{1 + 2\gamma_i - \Gamma_{-i}}{2 - 2\gamma_i + \Gamma_{-i}} c_{-i},$$

where γ_i is the individual MPCR of player i, and $\Gamma_{-i} = \sum_{j \neq i} \gamma_j$.

The solution to maximization problem $V_i(c_i, c_{-i}) \to \max\limits_{c_i}$ is given by

$$c_i^*(c_{-i}) = \begin{cases} m^{k(i)}(c_{-i}) - \frac{1-\gamma_i}{\beta_i}, & \beta_i > 0, \\ 0, & \beta = 0. \end{cases} \tag{I.22}$$

Note that the optimal contribution in (I.22) implies that contributions should be *lower* than the fairness ideal suggests. This implies that neither equilibria with small contributions nor equilibria in which all players contribute full endowments can exist with these preferences. I derive equilibrium contributions below.

Similar to Neitzel and Sääksvuori (2013), one can study what fairness ideals are preferred by players with different returns from the public good. The utility function in

[38] This utility function was first introduced in Cappelen et al. (2007).

the optimum takes values

$$V_i(c_i^*(c_{-i}), c_{-i}) = y + \gamma_i c_{-i} - (1-\gamma_i) m^{k(i)}(c_{-i}) + \frac{(1-\gamma_i)^2}{2\beta_i}.$$

For given contributions of others, c_{-i}, the utility function (I.21) takes maximum value for those fairness ideas that maximize the term $\gamma_i c_{-i} - (1-\gamma_i) m^{k(i)}(c_{-i})$. Table I.B.1 below shows the values of the term $\gamma_i - (1-\gamma_i) m^{k(i)}(c_{-i})/c_{-i}$ for the players with different returns from the public good in different groups for the parameter values in my experiment.

Table I.B.1: Values of $\gamma_i - (1-\gamma_i) m^{k(i)}(c_{-i})/c_{-i}, i = \{L, H\}, \gamma_L = \frac{1}{2}, \gamma_H = \frac{3}{4}$ for different fairness ideals $k(i)$

		Equality	Equity
H-HL groups	L-player	1/4	2/5
	H-player	5/8	4/7
L-HL groups	H-player	5/8	1/2
	L-player	1/4	1/3

According to the table, both in H-HL and in L-HL groups L-players prefer equal payoffs, while H-players prefer equal contributions. This prediction is similar to the one obtained in Proposition 1 in Neitzel and Sääksvuori (2013) for heterogeneously-endowed players, and goes in line with the results obtained in Reuben and Riedl (2013).

Proposition 7. *In a simultaneous contribution game, in both types of groups, players with high returns from the public good prefer equal contributions (equality), while players with low returns from the public good prefer equal payoffs (equity).*

This brings an idea what to expect in a sequential game. If the H-leader contributes first, then one can expect the H-follower to match the leader's contribution, because H-players prefer equality over other fairness ideals, while the L-follower, who prefers equity over other fairness ideals, will match the leader's payoff and thus contribute less than the leader. This situation corresponds to the equality-of-payoffs normatively appealing contribution rule.

At the same time, if the L-leader contributes first, then one can expect equal contributions from both followers: from the L-follower because she prefers equity, and by matching the leader's contribution she also matches the leader's payoff; and from the H-follower because she prefers equality and matches the leader's contribution. This situation corresponds to the equality-of-contributions normatively appealing contribution rule.

Proposition 8. *In a sequential game, players in H-HL groups contribute according to the equality-of-payoffs contribution rule, and players in L-HL groups contribute according to the equality-of-contributions rule.*

Equilibrium Contributions

Without loss of generality, let $m^{k(i)} = A^{k(i)} c_{-i}$. Then one can substitute the optimal solution for player j, $c_j^*(c_i, c_l)$, into (I.22) (consider the non-trivial case $\beta_i > 0$):

$$c_i(c_l) = A^{k(i)} \left(\underbrace{A^{k(j)} (c_i(c_l) + c_l) - \frac{1-\gamma_j}{\beta_j}}_{c_j^*(c_i, c_l)} + c_l \right) - \frac{1-\gamma_i}{\beta_i},$$

which gives us

$$c_i^*(c_l) = \left(A^{k(i)} \left[\left(A^{k(j)} + 1 \right) c_l - \frac{1-\gamma_j}{\beta_j} \right] - \frac{1-\gamma_i}{\beta_i} \right) \Big/ \left(1 - A^{k(i)} A^{k(j)} \right). \tag{I.23}$$

The solution given in (I.23) corresponds to the optimal contribution of the follower in a sequential game, given the contribution of the leader she observes, c_l. For the optimal contribution of the leader in a sequential game, as well as for the optimal contribution in a simultaneous game, one needs to substitute $c_l^*(c_i)$ into (I.23). This leads to

$$c_i^* = \frac{\frac{1-\gamma_i}{\beta_i} + \frac{A^{k(i)}(A^{k(j)}+1)}{1-A^{k(l)}A^{k(j)}} \frac{1-\gamma_l}{\beta_l} + \left[A^{k(i)} + \frac{A^{k(i)}A^{k(l)}(A^{k(j)}+1)}{1-A^{k(l)}A^{k(j)}} \right] \frac{1-\gamma_j}{\beta_j}}{\frac{A^{k(i)}A^{k(l)}(A^{k(j)}+1)^2}{(1-A^{k(p)}A^{k(j)})(1-A^{k(i)}A^{k(j)})} - 1}. \tag{I.24}$$

I.C Individual Patterns of Contributions and Beliefs

I.C.1 Individual Conditional Contributions

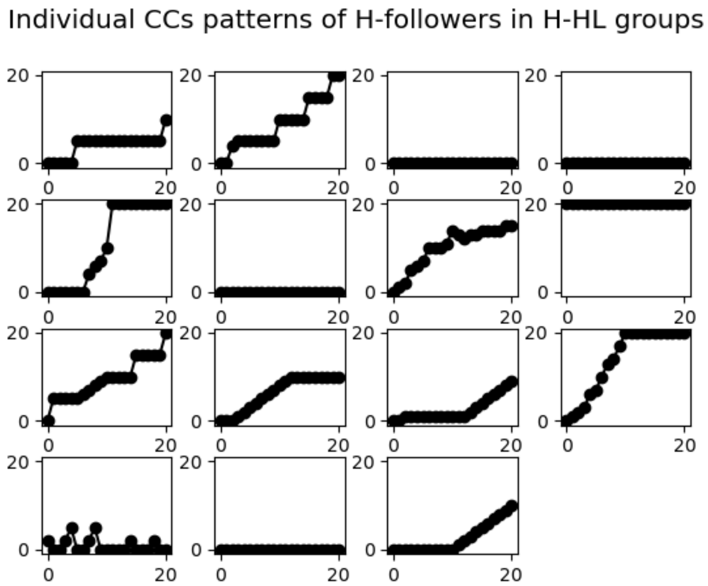


Figure I.C.1: Conditional contributions of H-followers in **H-HL groups** (exogenous returns)

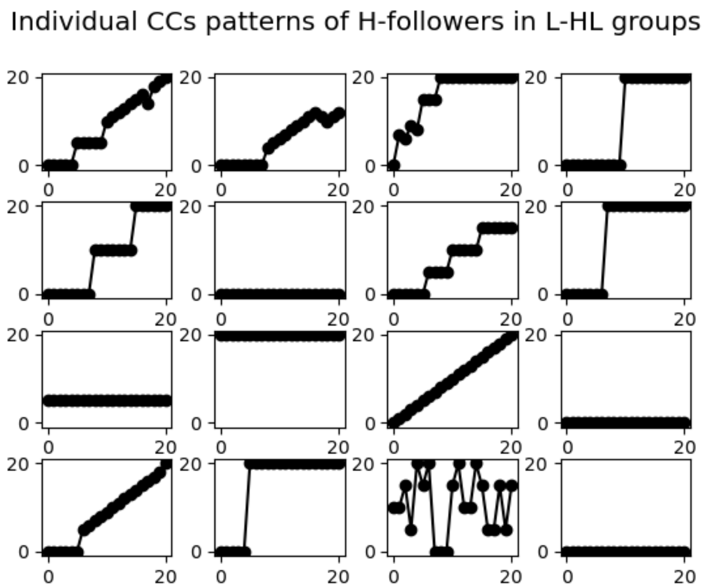


Figure I.C.2: Conditional contributions of H-followers in **L-HL groups** (exogenous returns)

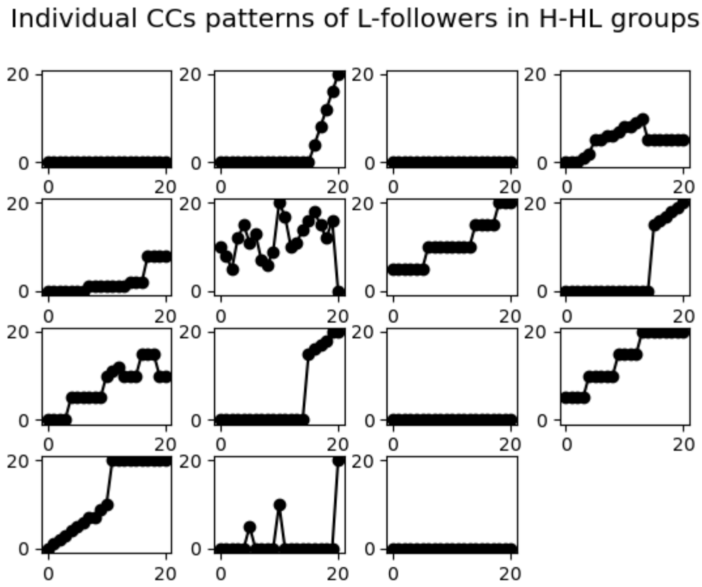


Figure I.C.3: Conditional contributions of L-followers in **H-HL groups** (exogenous returns)

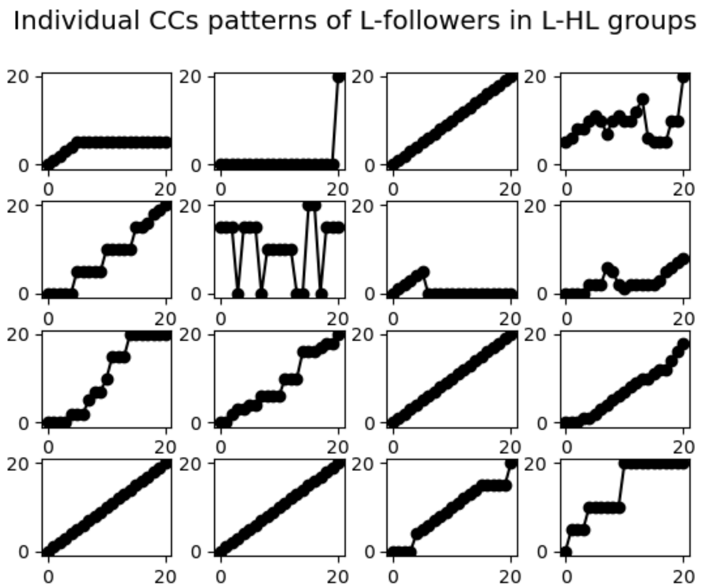


Figure I.C.4: Conditional contributions of L-followers in **L-HL groups** (exogenous returns)

I.C.2 Individual Conditional Beliefs

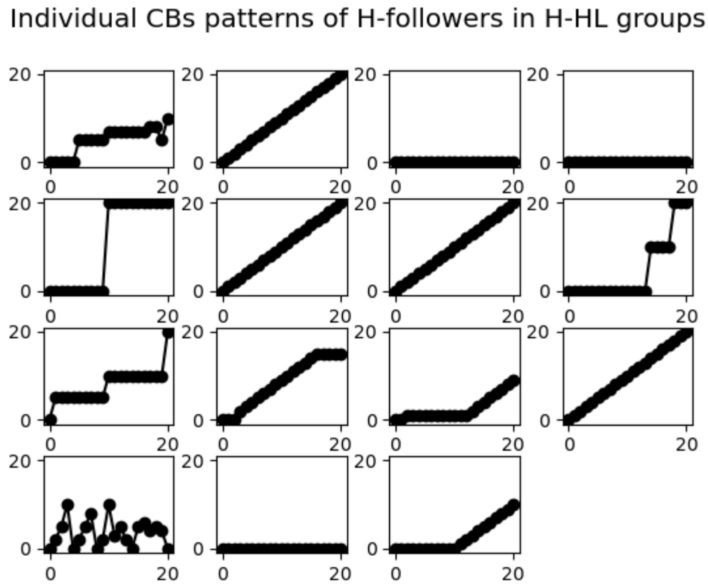


Figure I.C.5: Conditional beliefs of H-followers in **H-HL groups** (exogenous returns)

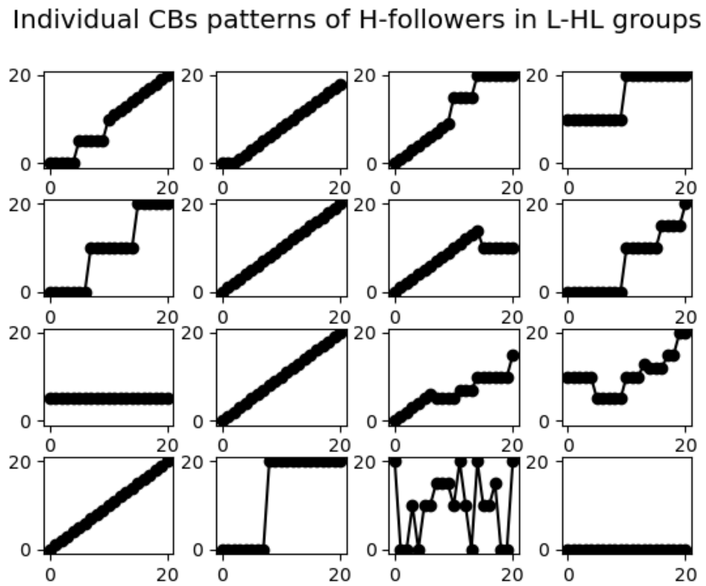


Figure I.C.6: Conditional beliefs of H-followers in **L-HL groups** (exogenous returns)

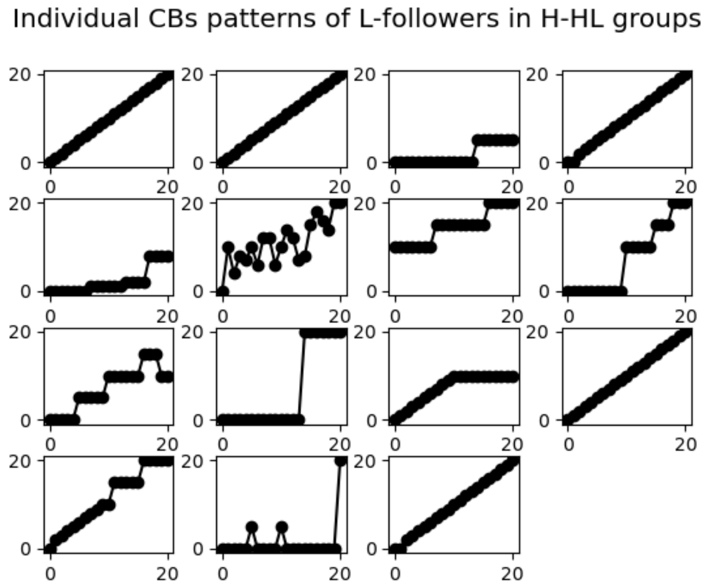


Figure I.C.7: Conditional beliefs of L-followers in **H-HL groups** (exogenous returns)

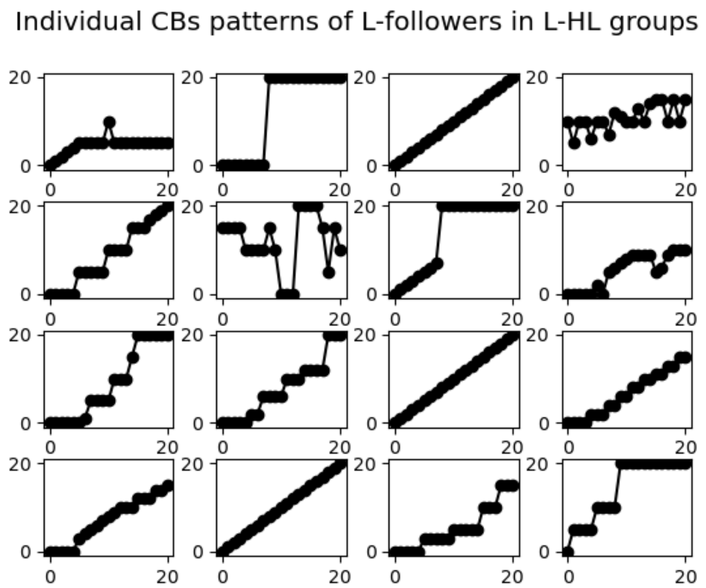


Figure I.C.8: Conditional beliefs of L-followers in **L-HL groups** (exogenous returns)

I.D Additional Regression Results

I.D.1 Additional Results for a Random Effects Model for Contributions in a One-Shot Game

Table I.D.1 shows estimates for a similar model estimated for each group separately: columns 1–3 were estimated on the sub-sample of H-HL groups, while columns 4–6 were estimated on the sub-sample of L-HL groups. Models in columns 1 and 4 use only the leader's contribution, dummies for the follower's low return and cross-terms as explanatory variables; models in columns 2 and 5 add beliefs about (conditional) contribution of another follower in the group; and models in columns 3 and 6 also include individual controls. All standard errors are clustered at the individual (participant's) level.

Table I.D.2 shows the comparison of contribution patterns between groups of followers with different returns. Columns 1 to 3 correspond with the estimates on the sub-sample of H-followers, while columns 4 to 6 correspond with the estimates on the sub-sample of L-followers. As in the previous tables, columns 1 and 4 show results for the simplest model (without beliefs and additional controls), models in columns 2 and 5 add beliefs and corresponding cross-terms as explanatory variables, and columns 3 and 6 include additional individual controls from the questionnaire.

The results presented in Table I.D.1 indicate differences between followers' contribution patterns within groups: in H-HL groups, H-followers contribute on average more than L-followers once individual controls are included. Moreover, a between-group comparison confirms the heterogeneous effect of the leader's contribution that was shown in Table I.3, namely that in H-HL groups (with H-leaders) an increased leader's contribution results in a very small increase in followers' contributions compared with L-HL groups: while an increase in the leader's contribution by one point increases (on average and given the rest equal) followers' contributions by around 0.33 points in L-HL groups, the same increase in the leader's contribution in H-HL groups only results in roughly a 0.1 increase in points contributed by followers.

The results presented in Table I.D.2 show substantial differences in contribution patterns of followers with identical returns between groups. First, once individual controls are included (columns 3 and 6), it appears that both H- and L-followers contribute significantly less in H-HL groups than in L-HL groups given the same contribution of the leader (and the rest being equal). This goes in line with behavioral predictions: not only do followers contribute less on average in H-HL groups, but they also increase their contributions in response to increased contribution of the leader less in H-HL groups than in L-HL groups. Moreover, the effect of being led by H-leader eliminates the effect of the leader's contribution on the decision of the followers (Wald tests for the hypothesis that coefficients for the leader's contributions and for cross-term H-leader × leader's contributions sum up to zero: $p = 0.075$ for H-followers and $p > 0.5$ for L-followers).

Table I.D.1: Estimates for a random effects model of conditional contributions by group (exogenous returns)

	(1) H-HL	(2) H-HL	(3) H-HL	(4) L-HL	(5) L-HL	(6) L-HL
Leader's contribution	0.441***	0.089	0.092	0.667***	0.321*	0.318*
	(0.118)	(0.050)	(0.050)	(0.138)	(0.125)	(0.125)
L-follower	-1.392	-1.661	-1.593	-1.043	-0.703	-1.987
	(1.692)	(1.644)	(1.556)	(1.803)	(1.749)	(2.149)
L-follower × Leader's contribution	0.056	-0.132	-0.135	0.020	0.144	0.143
	(0.163)	(0.093)	(0.094)	(0.180)	(0.176)	(0.176)
Belief		0.576***	0.570***		0.450***	0.454***
		(0.157)	(0.157)		(0.124)	(0.126)
L-follower × Belief		0.100	0.106		-0.187	-0.186
		(0.191)	(0.191)		(0.213)	(0.214)
male			4.690**			1.724
			(1.581)			(1.503)
Thinks that generally avoids risk			2.140			-1.149
			(1.836)			(1.652)
Thinks that generally accepts risk			0.839			3.676
			(2.139)			(2.412)
Agrees: can trust people in general			-6.790			-2.775
			(3.948)			(3.941)
Agrees: can rely on others			-5.660			-3.420
			(3.840)			(4.008)
Agrees: can trust strangers			3.214			4.560*
			(4.463)			(2.088)
Constant	1.599	1.834	1.252	2.122	1.563	-0.057
	(1.364)	(1.472)	(4.096)	(1.563)	(1.632)	(4.230)
Observations	630	630	630	672	672	672
Number of individuals	30	30	30	32	32	32

Notes: *L-follower* is a dummy equal to 1 if the follower has a low return from the public good; *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader; Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

Table I.D.2: Estimates for a random effects model of conditional contributions by follower (exogenous returns)

	(1) H-followers	(2) H-followers	(3) H-followers	(4) L-followers	(5) L-followers	(6) L-followers
Leader's contribution	0.667***	0.323**	0.320*	0.687***	0.465***	0.457***
	(0.138)	(0.124)	(0.125)	(0.115)	(0.124)	(0.123)
H-leader	-0.523	0.269	-0.778	-0.871	-0.686	-1.766
	(2.075)	(2.199)	(2.132)	(1.345)	(0.965)	(1.257)
H-leader × Leader's contribution	-0.226	-0.235	-0.235	-0.190	-0.510***	-0.508***
	(0.182)	(0.134)	(0.133)	(0.160)	(0.146)	(0.147)
Belief		0.447***	0.452***		0.263	0.273
		(0.125)	(0.125)		(0.173)	(0.171)
H-leader × Belief		0.129	0.130		0.414*	0.413*
		(0.200)	(0.201)		(0.204)	(0.204)
Constant	2.122	1.566	-0.419	1.079	0.860	2.068
	(1.564)	(1.635)	(2.549)	(0.900)	(0.628)	(1.855)
Other controls	No	No	Yes	No	No	Yes
Observations	651	651	651	651	651	651
Number of individuals	31	31	31	31	31	31

Notes: *H-leader* is a dummy equal to 1 if the leader in the group was selected to have a high return from the public good; *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader; Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

Second, once individual controls are included, L-followers rely on their beliefs about another follower's contribution more in H-HL groups than in L-HL groups, while there is no such difference for H-followers.

I.D.2 Additional Results for a Fixed Effects Model for Contributions in a One-Shot Game

Table I.D.3: Estimates for a fixed effects model of conditional contributions (exogenous returns)

	(1) Contribution	(2) Contribution
Leader's contribution	0.667***	0.327**
	(0.137)	(0.123)
L-follower × Leader's contribution	0.020	0.136
	(0.178)	(0.171)
H-leader × Leader's contribution	-0.226	-0.234
	(0.180)	(0.132)
L-follower × H-leader × Leader's contribution	0.036	-0.271
	(0.240)	(0.196)
Belief		0.443***
		(0.124)
H-leader × Belief		0.126
		(0.199)
L-follower × Belief		-0.177
		(0.208)
L-follower × H-leader × Belief		0.282
		(0.282)
Constant	1.263*	1.112*
	(0.603)	(0.520)
Observations	1302	1302
Number of individuals	62	62
Fixed Effects	0.46	0.56
model	fe	fe

Notes: *L-follower* is a dummy equal to 1 if the follower has a low return from the public good; *H-leader* is a dummy equal to 1 if the leader in the group was selected to have a high return from the public good; *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader.

Table I.D.4: Estimates for a fixed effects model of conditional contributions by group (exogenous returns)

	(1) H-HL	(2) H-HL	(3) L-HL	(4) L-HL
Leader's contribution	0.441***	0.093	0.667***	0.327*
	(0.118)	(0.051)	(0.138)	(0.123)
L-follower × Leader's contribution	0.056	-0.135	0.020	0.136
	(0.163)	(0.095)	(0.179)	(0.173)
Belief		0.569**		0.443**
		(0.157)		(0.125)
L-follower × Belief		0.106		-0.177
		(0.192)		(0.210)
Constant	0.903	1.003	1.600	1.214
	(0.813)	(0.621)	(0.897)	(0.832)
Observations	630	630	672	672
Number of individuals	30	30	32	32
Fixed Effects	0.39	0.57	0.50	0.55
model	fe	fe	fe	fe

Notes: *L-follower* is a dummy equal to 1 if the follower has a low return from the public good; *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader.

Table I.D.5: Estimates for a fixed effects model of conditional contributions by follower (exogenous returns)

	(1) H-followers	(2) H-followers	(3) L-followers	(4) L-followers
Leader's contribution	0.667***	0.327*	0.687***	0.463***
	(0.138)	(0.124)	(0.115)	(0.121)
H-leader × Leader's contribution	-0.226	-0.234	-0.190	-0.505**
	(0.182)	(0.134)	(0.160)	(0.145)
Belief		0.443**		0.266
		(0.125)		(0.169)
H-leader × Belief		0.126		0.408
		(0.201)		(0.202)
Constant	1.869*	1.697*	0.657	0.526
	(0.913)	(0.742)	(0.803)	(0.741)
Observations	651	651	651	651
Number of individuals	31	31	31	31
Fixed Effects	0.43	0.55	0.48	0.57
model	fe	fe	fe	fe

Notes: *H-leader* is a dummy equal to 1 if the leader in the group was selected to have a high return from the public good; *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader.

Table I.D.6: Estimates for fixed effects model of conditional contributions with respect to contributions by the leader with similar / different return (exogenous returns)

	(1) H-followers	(2) L-followers
Contribution by leader with same return	0.093	0.463***
	(0.051)	(0.121)
Contribution by leader with different return	0.327*	-0.042
	(0.124)	(0.081)
Belief, when leader has same return	0.569**	0.266
	(0.157)	(0.169)
Belief, when leader has different return	0.443**	0.675***
	(0.125)	(0.111)
Constant	1.697*	0.526
	(0.742)	(0.741)
Observations	651	651
Number of individuals	31	31
Fixed Effects	0.55	0.57
model	fe	fe

Notes: *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader.

I.E Summary Statistics for Repeated Games

Tables I.E.1 and I.E.2 below show basic summary statistics (mean, standard deviation, and median) on actual contributions and payoffs in repeated games (simultaneous and sequential, correspondingly), separately for periods 1, 2–5, 6–9, 10, and over the whole game.

Table I.E.1: Summary statistics for the outcomes in a repeated simultaneous game (exogenous returns)

Player	**Stat**	**Periods**				**Overall**
		1	**2–5**	**6–9**	**10**	
Contributions						
	mean	10.4	8.25	5.19	2.75	6.69
HHL, L-player	std	(8.26)	(7.99)	(7.01)	(5.12)	(7.59)
	median	[11.5]	[5]	[0]	[0]	[5]
	mean	12.8	12.3	9.86	5.56	10.7
HHL, H-players	std	(6.31)	(7.49)	(8.49)	(7.85)	(8.05)
	median	[15]	[15]	[10]	[0]	[10]
	mean	14.7	14.3	11.3	7.56	12.4
LHL, L-players	std	(7.63)	(7.96)	(8.73)	(8.78)	(8.54)
	median	[20]	[20]	[10]	[2.5]	[16.5]
	mean	16.9	13.8	13.4	10.6	13.6
LHL, H-player	std	(7.04)	(9.16)	(8.27)	(9.43)	(8.6)
	median	[20]	[20]	[20]	[12.5]	[20]
Payoffs						
	mean	27.6	28.2	27.3	24.2	27.4
HHL, L-player	std	(5.96)	(4.9)	(5.56)	(5.26)	(5.33)
	median	[28.8]	[28.8]	[29.2]	[22.2]	[28.5]
	mean	34.2	32.3	28.8	24.8	30.4
HHL, H-players	std	(7.92)	(9.25)	(9.25)	(7.78)	(9.3)
	median	[35.8]	[31.8]	[26]	[20]	[29.9]
	mean	28.4	26.9	26.7	25.3	26.8
LHL, L-players	std	(7.47)	(5.48)	(5.44)	(6.79)	(5.81)
	median	[30]	[30]	[30]	[25]	[30]
	mean	37.8	38	33.5	28.7	35.2
LHL, H-player	std	(5.46)	(9.22)	(10.5)	(9.75)	(9.85)
	median	[37.5]	[41.2]	[32.2]	[27.2]	[37.5]

Table I.E.2: Summary statistics for the outcomes in a repeated sequential game (exogenous returns)

Player	Stat	Periods 1	2–5	6–9	10	Overall
Contributions						
H-HL, leader	mean	10.7	10.8	11.8	5	10.6
	std	(8.38)	(7.94)	(8.46)	(7.64)	(8.23)
	median	[10]	[10]	[12.5]	[0]	[10]
H-HL, L-follower	mean	12.1	8.32	8.07	2.29	8
	std	(9.06)	(8.76)	(8.25)	(4.27)	(8.4)
	median	[15]	[4]	[4.5]	[0]	[4.5]
H-HL, H-follower	mean	13.4	11.8	10.4	0.714	10.3
	std	(8)	(8.1)	(8.9)	(1.89)	(8.59)
	median	[15]	[13]	[9]	[0]	[10]
L-HL, leader	mean	18.1	14.8	15.2	9.38	14.7
	std	(3.72)	(6.57)	(6.63)	(10.2)	(6.98)
	median	[20]	[20]	[20]	[7.5]	[20]
L-HL, L-follower	mean	11.9	13.2	13.4	9.88	12.8
	std	(7.53)	(8.27)	(8.19)	(9.72)	(8.23)
	median	[10]	[20]	[20]	[9.5]	[20]
L-HL, H-follower	mean	18.1	16.2	15.4	5	14.9
	std	(3.72)	(6.31)	(7.27)	(9.26)	(7.55)
	median	[20]	[20]	[20]	[0]	[20]
Payoffs						
H-HL, leader	mean	36.5	32.4	30.9	21	31.1
	std	(8.45)	(8.72)	(10.1)	(4.68)	(9.59)
	median	[36.2]	[31]	[29.4]	[20]	[30.1]
H-HL, L-follower	mean	26	27.1	27.1	21.7	26.5
	std	(4.56)	(6.57)	(5.83)	(4.85)	(6.06)
	median	[27.5]	[27.8]	[28.2]	[20]	[27.5]
H-HL, H-follower	mean	33.8	31.3	32.3	25.3	31.4
	std	(12.4)	(10.1)	(9.97)	(5.67)	(10)
	median	[35.2]	[30]	[32]	[26.2]	[30]
L-HL, leader	mean	25.9	27.3	26.8	22.8	26.5
	std	(3.52)	(4.03)	(5.01)	(4.53)	(4.57)
	median	[25]	[30]	[30]	[20]	[30]
L-HL, L-follower	mean	32.2	28.8	28.6	22.2	28.4
	std	(4.32)	(3.4)	(3.71)	(4.83)	(4.36)
	median	[31.2]	[30]	[30]	[20]	[30]
L-HL, H-follower	mean	38	37	37.6	33.2	36.9
	std	(6.91)	(9.63)	(9.21)	(13.5)	(9.57)
	median	[38.1]	[45]	[45]	[32.8]	[45]

I.F Dynamics of Contributions within Groups

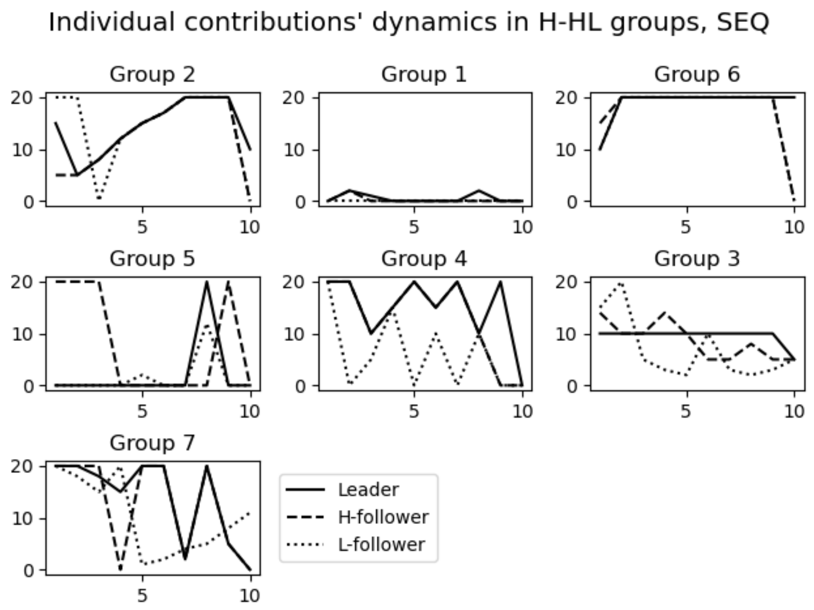


Figure I.F.1: Dynamics of contributions in **H-HL groups**, sequential game (exogenous returns)

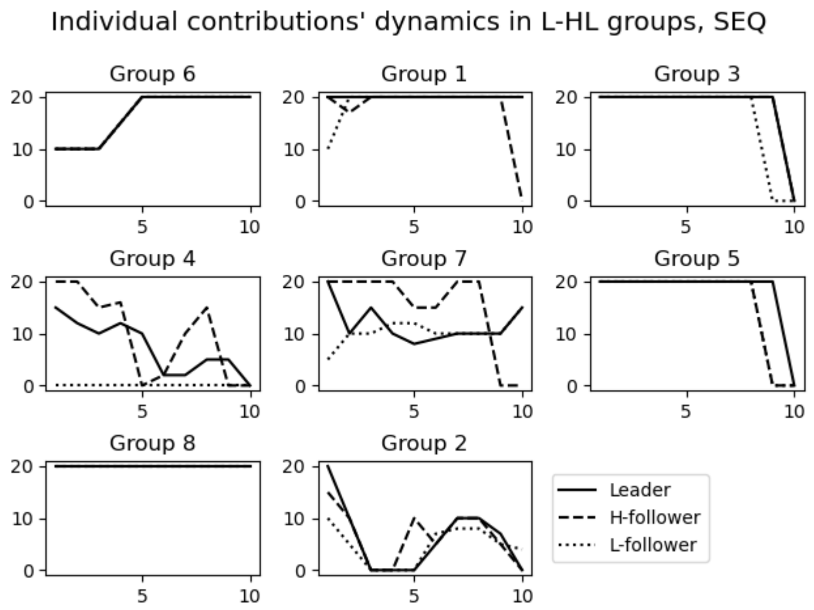


Figure I.F.2: Dynamics of contributions in **L-HL groups**, sequential game (exogenous returns)

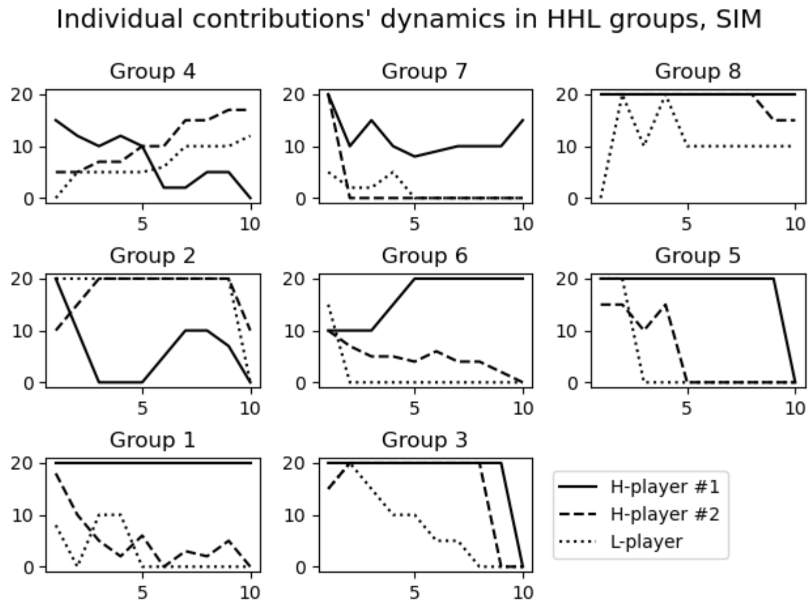


Figure I.F.3: Dynamics of contributions in **HHL groups**, simultaneous game (exogenous returns)

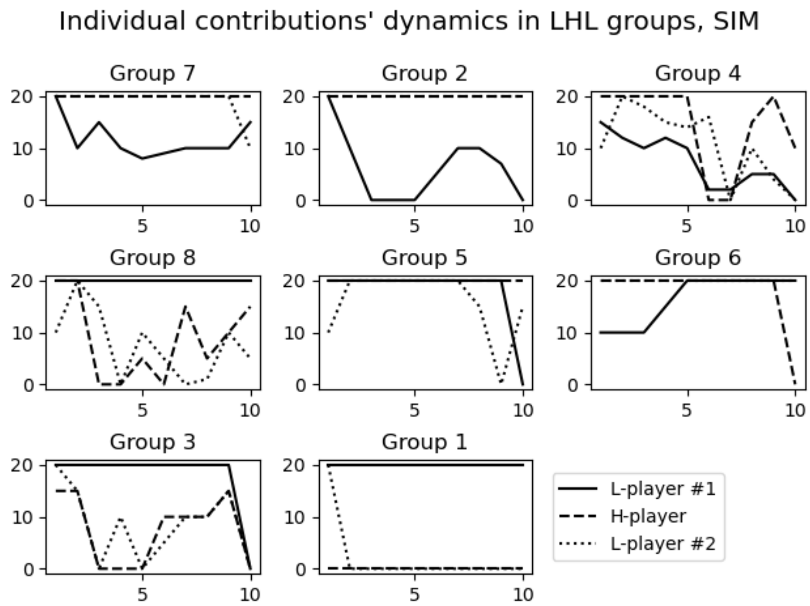


Figure I.F.4: Dynamics of contributions in **LHL groups**, simultaneous game (exogenous returns)

Chapter II

Endogenous Composition of Returns and Leading-by-Example in an Asymmetric Public Goods Game*

II.1 Introduction

When individuals are engaged in group activities, it is often difficult to assume that these individuals are identical. The presence of heterogeneity in returns from the public good increases tensions between contribution allocations that different players view as 'fair' ones (see Nikiforakis et al., 2012, 2015; Reuben and Riedl, 2013). People may base their decisions on asymmetries in endowments or returns (see Fischbacher et al., 2014; Glöckner et al., 2011; Levati et al., 2007). Moreover, the effectiveness of the institutions meant to foster cooperation may also depend on such asymmetries (see, for example, Kube et al., 2015). Leadership has shown to be an institution that improves cooperation (see Frackenpohl et al., 2016; Gächter et al., 2012; Güth et al., 2007), as it serves as a coordination mechanism.[1]

In my previous chapter, I show that, in a sequential public goods game with asymmetric returns, followers are likely to cooperate with a leader who benefits relatively less from the public good than the one who benefits relatively more. There are two potential explanations for this behavior. One is that the contribution of such a leader is considered to be a "clear" signal about the suggested contribution norm.[2] The second is that low-benefiting leaders are viewed as "most disadvantaged" players in the group, and their contribution is considered to be a "risky investment" in the group's favor. It seems commonly accepted that being a successful leader usually calls for potential self-sacrifices ("no pain, no gain", see p. 3 in Jensen, 2005). In the literature on leadership in social dilemmas, economists have suggested several potential motives that may explain why leaders willingly accept to

*I would like to thank Sebastian Kube and Sebastian Schaube for their insightful comments. Financial support from the Institute of Applied Microeconomics (IAME) is gratefully acknowledged.

[1]In the simplest case with just the sequential order of moves (called leading-by-example), one person makes a decision first (the leader), the others (the followers) observe this decision and make their decisions afterwards.

[2]High contributions by a low-benefiting leader excludes the possibility that the leader suggests a contribution rule that results in equal payoffs. At the same time, high contribution made by a high-benefiting leader does not clearly imply whether she suggests to contribute equally or achieve equal payoffs.

put their self-interests at risk to guide their groups towards a desired collective goal. One of the central motives for this is the leader's expectations about the followers' reciprocity to the contribution made by the leader (Arbak and Villeval, 2013; Gächter et al., 2010; Glöckner et al., 2011; Potters et al., 2001). However, it is unclear how leaders' expectations about reciprocal reactions and how the reactions themselves change if some players have the possibility to influence the asymmetries in returns among group members.

To shed light on these questions, in this paper I use an economic experiment to study contribution decisions in a sequential public goods game with asymmetric returns, where second-movers (followers) are allowed to choose the composition of returns in their group by voting on what return the leader (the first-mover) in their group should have, high or low.

My findings suggest that followers reciprocate the leaders' "sacrifice", albeit in different ways depending on their own return from the public good. Followers with high returns vote for the leader to have a high return and follow the actual leader's contribution quite closely in a one-shot game. By contrast, followers with low returns vote for the leader to have a high return as well but start free-riding more. The regression analysis shows that these differences in contribution responses are associated with players' own voting decisions: followers with low returns who vote for the leader to have a low return contribute quite generously regardless of the resulting leader's return. It seems that followers who are able to choose how much the first-mover benefits from the public good try to motivate their leaders to contribute larger amounts by voting for the leader to have high returns. Whereas followers with high returns may vote for such a composition of returns because they might expect higher profits from cooperation with a high-benefiting leader than with a low-benefiting one, followers with low returns seem to do so only to benefit from increased leaders' contributions. At the same time leaders are willing to contribute more if they receive high returns. This effect most likely originates from the reciprocation of the fact that (at least some) followers voted for the leader to have a high return and not from expecting higher contributions from the followers.

The rest of the paper is organized as follows. Subsequently, in Section II.2 I discuss the literature guiding this research. I present the design of the experiment in Section II.3 and formulate main hypotheses on what behavioral patterns one might expect in my experiment in Section II.4. The results are presented in Section II.5. Section II.6 presents the results from an additional treatment with a repeated sequential game. Finally, Section II.7 concludes.

II.2 Previous Research

This paper contributes to the existing literature on leadership in public goods games in two ways. First, it adds to the literature on public goods games with heterogeneous returns, as I allow the composition of returns in the group to change by varying the leader's return from the public good. Second, it adds to the literature on the effect democratic elections in public goods games, as I let followers choose the composition of returns in the group by voting on what return the leader in the group should have, high or low.

II.2.1 Literature on Public Goods Games with Asymmetric Returns

Normative conflicts may arise due to the co-existence of multiple plausible contribution rules: there might be no uniform perceptions on whether players should contribute equally, efficiently or make contributions that lead to equal payoffs.

Several papers establish the presence of normative conflicts when asymmetric payoffs are induced by making players differ in their benefits from the public good (see Fischbacher et al., 2014; Fisher et al., 1995; Reuben and Riedl, 2013; Van Dijk and Wilke, 1993). These conflicts may even start a feud if counter-punishment is allowed (see Nikiforakis et al., 2012). This feud offsets the efficiency gains from increased cooperation induced by the threat of being punished. Kube et al. (2015) show that such conflicts can prevent the formation of efficient institutions meant to foster cooperation.

II.2.2 Leadership and Endogenous Selection

The effectiveness of leadership institutions may strongly depend on the way in which leaders are appointed: whether they are selected exogenously, elected by other group members or even self-selected. Several papers suggest that contributions should be higher in sequential games with endogenously selected leaders compared with those with exogenously-appointed ones (see Güth et al., 2007; Haigner and Wakolbinger, 2010; Potters et al., 2005; Rivas and Sutter, 2011) at least in *symmetric groups*.[3] Endogenous leadership institutions that do not impose sequential move structure, for example, with democratic delegation of collective contribution decisions to a single member, might improve the provision of public goods as well (see Hamman et al., 2011).

Nevertheless, endogeneity of the institution does not always improve cooperation. Arbak and Villeval (2013) find that, despite having a self-selected first-mover in a repeated three-player public goods game improves followers' contributions, followers are more responsive to randomly-chosen leaders (selected each period). Centorrino et al. (2013) study the groups in which players have to compete for the first-mover position by bidding in a second-price auction prior to playing a repeated game. They find that such competition drives an increase in leaders' contributions, but has an ambiguous effect on the total provision of public good. Similar findings are reported in Ibanez and Schaffland (2018), who show that there is no difference in cooperation between groups with in-group leaders versus groups with out-group leaders.

Although I do not allow the players to be a member of a directly selected particular kind of group, the literature on this topic might also be relevant for my setup, as it reports

[3] Levati et al. (2007) also document the presence of this effect in groups with asymmetric endowments; however, their results might be driven by leaders' strategic motives.

differences in cooperation levels between self-selected groups.[4] For example, Brekke et al. (2011) find that players selecting by themselves to be in groups whose members donate extra payoff to the Red Cross sustain higher levels of cooperation throughout the experiment than players in groups whose members receive this extra payoff for themselves. My setup is different, because, first, I do not allow for the actual selection of groups, second, players in my setup play a sequential linear public goods game, and, finally, players choose the composition of returns in their groups (by voting on the return of the leader) rather than selecting the groups based on performance (Ehrhart and Keser, 1999) or institutional design (Gürerk et al., 2006).

Another relevant concept for asymmetric groups in my setup is group identity: players with different returns might see other players with the same returns as in-group members and players with returns that differ from their own return as out-group members. Drouvelis and Nosenzo (2013) find that common identity between the leader and followers enhances cooperation compared with treatments with no induced identity, but no difference is found when only part of the followers share the leader's identity, or when followers share a common identity that differs from that of the leader. In Drouvelis and Nosenzo (2013), the identity between players comes from the minimal group paradigm, namely that players differ in nothing but the label attached to them (red or blue), while in my setup the identity can have a direct relation to payoffs via the difference in returns from the public good.

II.3 The Design of the Experiment

The experiment comprises of two parts. There are two stages in the first part. Players randomly form groups of $n = 3$ people, with randomly assigned roles: one leader (the first-mover) and two followers (the second-movers). Players' individual payoffs π_i are determined according to the formula

$$\pi_i = (20 - c_i) + \gamma_i \sum_{j=1}^{n} c_j, \quad i = 1, \ldots, n, \tag{II.1}$$

where c_i is the contribution of player i, and $\gamma_i \in \{\gamma_L, \gamma_H\}$ represents an individual's MPCR from the public good, with $\gamma_H = 0.75$ and $\gamma_L = 0.5$.

The followers always have different returns (allocated randomly), high and low, correspondingly. In each group, there is always one follower with a high return (H-follower) and one with a low return (L-follower). The returns of the followers are randomly assigned. In the first stage, followers vote on whether the leader will have high or low returns from the public good (with ties split randomly). Thus, if the leader has a high return then there are two players with high returns from the public good and one player with a low return (such groups are denoted as H-HL), and if the leader has a low return then there are two

[4] In exogenously set groups of "like-minded" people cooperation is also higher than in most cooperative randomly composed groups, see Gächter and Thöni (2005).

players with low returns and one with a high return (such groups are denoted as L-HL). In the second stage, players make their contribution decisions according to the following procedure (strategy method, see Selten, 1967). The leader states how much she would like to contribute if she has a high return, and how much if she has a low return (without knowing the outcome of the voting procedure). The followers decide on their own contributions in the form of a contribution table: they have to state their contribution decisions for *all* possible contributions by the leader, *conditional on the leader's return.* Accordingly, the followers make contribution decisions prior to learning the outcome of the voting stage, i.e. they have to fill out two contribution tables separately: for the cases when the leader has a high return and when the leader has a low return from the public good. Actual contributions and payoffs are then realized according to the choices after the leader's return is determined through the voting decisions.

In addition to contribution tables, followers were asked to fill out tables with respective *conditional beliefs*:[5] guesses on how much they expect the other follower to contribute for all possible contributions of the leader, again, separately for the cases when the leader has a high return and when the leader has a low return from the public good. Leaders were asked about their beliefs about the contributions of followers, conditional on the leader's return.

This strategy method approach allows for a clean comparison of contribution profiles, as it helps to isolate actual decisions from reputation effects or potential emotional spillovers due to a discrepancy of expectations and realized outcomes (for example, disappointment or even spiteful behavior if a leader expects to have a high return but is selected to have a low one). However, one cannot exclude the possibility of simultaneous decisions on voting and contribution: followers may vote strategically for the leader to have a high return to motivate their leader to contribute high amounts, but lower their own contributions with respect to the leader's decision.

In the second part of the experiment subjects play a repeated game for ten consecutive periods in the same groups as before. The payoffs in *each* period are determined according to equation (II.1). Prior to the game, followers vote on the return of the leader (ties split randomly). The leader with the selected return will contribute first in *all* ten periods of the game. Each period of the repeated game has two stages: first, the leader contributes, then, the followers observe the leader's actual decision and decide on their own contributions. After each period all group members learn the respective contributions and payoffs of other players.

In order to prevent any spillover effects from the outcomes in the first part on the decisions in the second part, the actual outcomes in a one-shot game are only reported at the end of the experiment, after the repeated game was played. The followers kept the same return from the public good throughout the experiment (same as in the one-shot game). All players had the same role in sequential repeated games as in the one-shot game (i.e. leaders in the one-shot game were leaders in the sequential repeated game, etc.).

[5] Following Fischbacher and Gächter (2010), I call players' expectations *players' beliefs*.

II.3.1 Procedures

The instructions (see Appendix II.A) explained to the participants the payment system in a public good problem and the differences between players. After participants had read the instructions, they had to answer control questions to ensure their understanding of the payoff structure and the dilemma situation. The experiment started only after all participants answered all questions correctly. I use neutral wording in the instructions and during the experiment (e.g. players with low and high returns are referred to as players of type A and of type B, respectively, with known MPCR).

At the end of the experiment, participants were asked to complete a questionnaire with questions on gender, age and the field of study, self-reported measures of risk preferences, trust and power distance, as well as the questions on the importance of different factors on their decisions (how important for their decisions were: contributions of others, own payoffs, payoffs of others, etc.).

The experiment took place at the BonnEconLab in February 2018. All treatments were programmed using the software z-Tree (Fischbacher, 2007). The participants were mostly students from various disciplines from the University of Bonn, recruited using HROOT recruitment software (Bock et al., 2014). There were two sessions, with 24 participants in each session. All 48 participants were randomly assigned into groups of three. Each group member was randomly assigned the role (a leader or a follower), and followers were randomly assigned their returns (high or low). In total, I had sixteen independent groups of three players each.

On average participants earned 30.3 EUR (around 37.7 USD) including a show-up fee of 4 EUR.[6] Subjects were paid up to 2 Euros for the correct guess.[7] Each session lasted 120 minutes (including random allocation of the participants into cubicles, quiz, questionnaire and payments).

II.4 Hypotheses

Introduction of social preferences can transform a social dilemma represented by a public goods game into a coordination problem with multiple equilibria. Given the multitude of social preference models, I try to formulate hypotheses under the assumption that players do take social comparison into account but without focusing on any specific form.

In the analysis below it is convenient to refer to contribution rules that players might follow: these contribution rules work as a coordination device (see the detailed discussion in Young, 2008, on social norms). Reuben and Riedl (2013) select several normatively ap-

[6] During the experiment subjects earned their payoffs in points, at an exchange rate of 1 point = 0.6 EUR for the points earned in a one-shot game, and at a rate of 1 point = 0.03 EUR for the points earned in a repeated game.

[7] Each participant was paid 1 EUR for making a correct guess. Although followers had to make 21 guesses, there was only one relevant *actual* contribution of the other follower, whereas for the leader there were two relevant *actual* contributions of the followers: in order to make these guesses incentive-compatible, I paid 1 EUR for each correct guess of the leader.

pealing contribution rules that can potentially arise in groups with heterogeneous returns that do not necessarily achieve efficiency in the provision of the public good: the equality of contributions and the equality of payoffs.

The equality of contributions implies that all players should contribute the same amount, independently of their individual MPCR. Given that the leader moves first and that all players should contribute similar amounts, their contributions should not depend on what return for the leader they vote on. The equality of payoffs suggests that players with low returns (L-players) should contribute less than players with high returns (H-players). Here the differences in voting decisions come to play. If players vote for the leader to have a low return, there will be two players with low returns and one player with a high return (denote it as L-HL group). In this case, in order to equalize payoffs, the L-leader and the L-follower contribute any amount between 0 and 10 and the H-follower should contribute double this amount to receive the same payoff. Instead, if players vote for the leader to have a high return, there will be two players with high returns and one player with a low one (denote it as L-HL group). Now, to equalize payoffs, H-leader and H-follower contribute any amount between 0 and 20, and L-follower should contribute 2/5 of that.

There are several important predictions that can be made from this perspective. First, choosing the return of the leader does not directly affect followers' payoffs. Second, the maximum payoff that followers can receive if they follow the equality-of-payoffs rule is 30 in L-HL groups, and 36 in H-HL groups. Instead, if they adhere to the equality-of-contributions rule, the maximum payoffs are now 45 for H-followers and 30 for L-followers. Third, note that in H-HL groups, H-players receive higher payoffs if they both contribute positive amounts than if they do not contribute at all (i.e. assuming that the H-follower does not free-ride), even if L-follower completely free-rides. Thus, if one assumes that all players adhere to the same contribution rule (induced by the first-mover), followers always vote for the leader to have a high return because their payoffs will be higher.[8] However, what voting behavior can be expected if H- and L-players prefer different contribution rules?

It is easy to show that the utility function with a quadratic loss from non-compliance to preferred contribution norm introduced in Cappelen et al. (2007) predicts the results obtained in Reuben and Riedl (2013) for groups with heterogeneous returns. Players with high returns prefer the equality-of-contributions rule and players with low returns prefer the equality-of-payoffs rule (derivation is similar to the way it is done in Neitzel and Sääksvuori, 2013).[9] If one assumes that players may adhere to different contribution rules, i.e. players with high returns prefer the equality-of-contributions rule and players with low returns prefer the equality-of-payoffs rule, one can show that the predictions about voting

[8] For the same level of the leader's contribution, x, followers in H-HL groups get $20+1.4x$, while in L-HL groups they get $20+x$ under the equality-of-payoffs rule. They are indifferent between L- and H-leaders under the equality-of-contributions rule.

[9] The intuition behind this choice is minimizing the disutility of the deviation from contributions implied by one of the rules in consideration (see also Appendix B.3 in Ananyev, 2019, for details).

behavior now depend on their expectations about the contribution of the leader. If players contribute according to these patterns, L-followers should always vote for the leader to have a high return since their expected payoffs are then higher.[10] H-followers should vote for H-leaders only if they expect L-leaders to contribute substantially less than H-leaders (otherwise, their expected payoffs are higher with L-leaders).[11]

One can expect the leader to contribute more if she receives a high return not only because it is profitable for her to cooperate with H-follower, but also reciprocating the voting decisions of the followers.

II.5 Results

In this section, I present the results of the analysis of the data obtained in the first part of the experiment.[12] In the one-shot game, contribution decisions are not affected by reputation concerns or emotional effects. I find that, despite most followers (both with high and low returns from the public good) voted for their leader to have a high return from the public good, followers with low returns do not reciprocate the leader's contribution at all, but condition their contribution decision solely on their own beliefs about the contribution of the other follower in the group.

I start with presenting summary statistics on the realized outcomes: voting decisions and realized contributions and payoffs. Table II.1 below shows summary statistics on voting decisions of the followers. Despite the predicted heterogeneity in voting decisions between H- and L-followers, more than 80% of all followers voted for their leader to have a high return from the public good.

Table II.1: Voting decisions of followers, %

	Voted	
Follower	**Low**	**High**
L-follower	12.5	87.5
H-follower	25	75

Table II.2 shows summary statistics on realized contributions and realized payoffs in the one-shot game. According to Table II.2, contributions of L-followers are almost three times lower than leaders' contributions in all groups, regardless of the leader's return

[10] Given the leader's contribution of x, the L-follower's payoff (under the patterns described above) is $20+0.5x$ in L-HL, and is $20+0.8x$ in H-HL (here, I assume that the followers take the leader's contribution as a reference). If H-leaders contribute more than L-leaders, L-followers are even more better off.

[11] Given the leader's contribution of x, the H-follower's payoff (under the patterns described above) is $20+1.25x$ in L-HL, and is $20+0.8x$ in H-HL (here, I assume that the followers take the leader's contribution as a reference). Therefore, for H-followers to prefer H-leaders, H-leaders should contribute 56.25% more than L-leaders.

[12] Computations for the analysis presented in this paper were programmed in Python (except for the regression analysis) using the Templates for Reproducible Research Projects in Economics by von Gaudecker (2014).

(Wilcoxon signed-rank tests: $p = 0.011$ for differences between contributions of leaders and L-followers in H-HL groups, and $p = 0.109$ for differences between contributions of leaders and L-followers in L-HL groups).[13] Contributions of H-followers also differ from contributions of leaders in H-HL groups (Wilcoxon signed-rank tests $p = 0.043$), but not in L-HL groups (Wilcoxon signed-rank tests $p = 1.0$). Regarding the payoffs, the only statistically significant difference is between payoffs of H-followers and H-leaders in H-HL groups (Wilcoxon signed-rank tests $p = 0.043$).

Table II.2: Summary statistics for the **realized** outcomes in the one-shot game (endogenous returns)

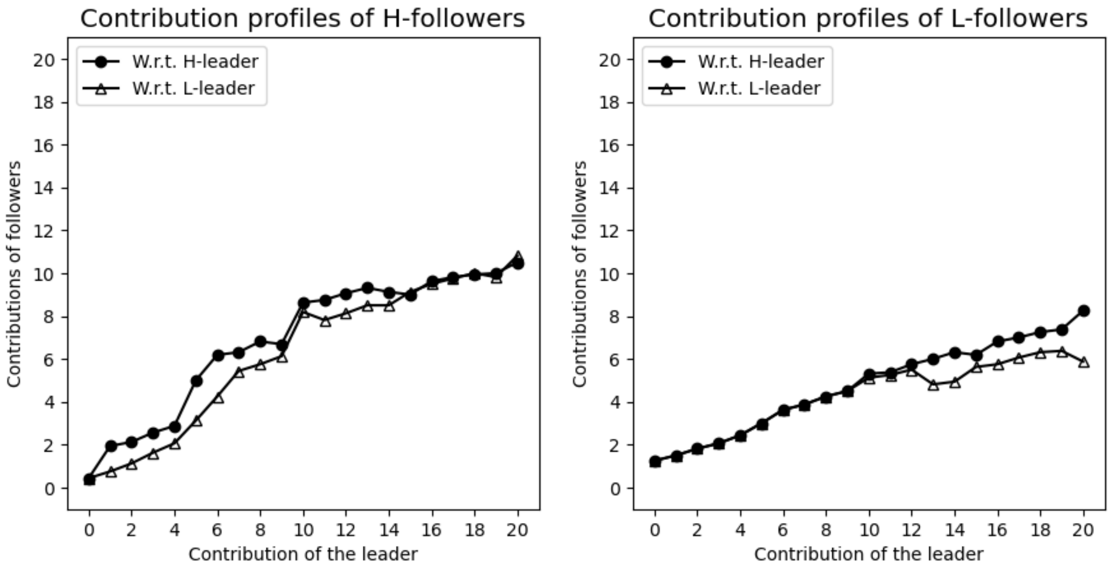
Player	**# obs**	**Mean**	**St.dev.**	**Min**	**Max**	**Median**
Contributions						
H-HL, leader	11	14.7	6.75	5	20	20
H-HL, L-follower	11	5.45	8.2	0	20	0
H-HL, H-follower	11	9.64	9.34	0	20	5
L-HL, leader	5	9	8.94	0	20	10
L-HL, L-follower	5	3.6	4.16	0	10	3
L-HL, H-follower	5	7.2	8.47	0	20	5
Payoffs						
H-HL, leader	11	27.6	10.5	16.5	45	25
H-HL, L-follower	11	29.5	6.57	20	40	30
H-HL, H-follower	11	32.7	7.66	22.5	45	34.2
L-HL, leader	5	20.9	4.25	14	25	22.5
L-HL, L-follower	5	26.3	5.76	20	35	26
L-HL, H-follower	5	27.6	8.21	18.8	37.5	28.5

The observed differences in contributions could occur due to various reasons, including contributions by leaders, differences in individual contribution preferences or beliefs about the contribution of the other follower. Below, I present conditional contributions of followers with respect to potential H- and L-leaders, correspondingly. Figure II.1 below shows within-subject differences in conditional contributions of H- and L-followers with respect to contributions of potential H- and L-leaders. Both H- and L-followers do not demonstrate differences in responses to the contributions of leaders with high or low returns. H-followers typically contribute more than L-followers, as Figure II.2 suggests.

Result 11. *Followers' contribution profiles do not differ between leaders with different returns.*

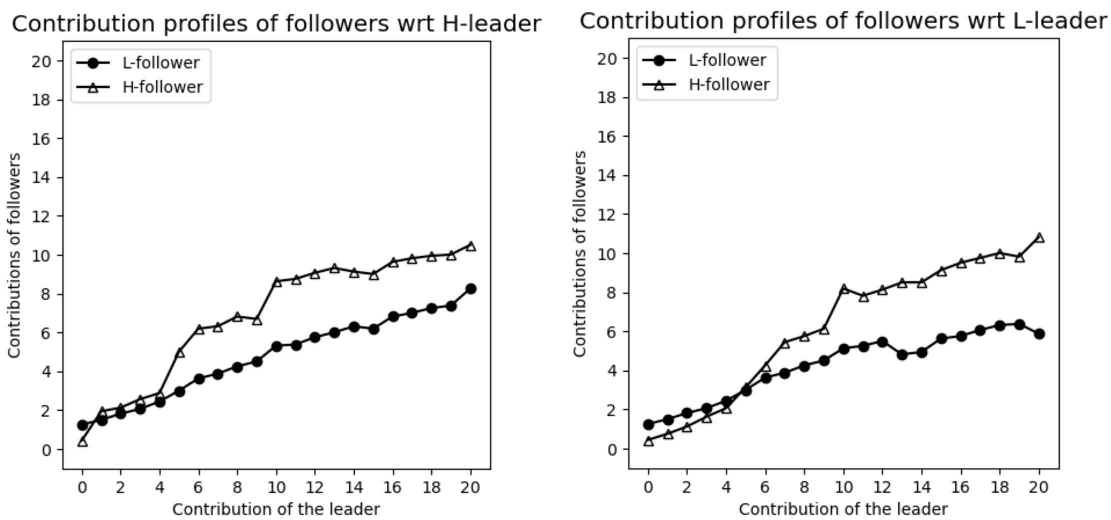
Result 12. *H-followers are willing to contribute more than L-followers in response to almost any contribution of the potential leader, both for H- and L-leaders.*

[13] All p-values are for two-sided tests, if not mentioned otherwise.



(a) H-followers (b) L-followers

Figure II.1: Within subject differences in conditional contributions with respect to *H*- and *L-leaders* (endogenous returns)

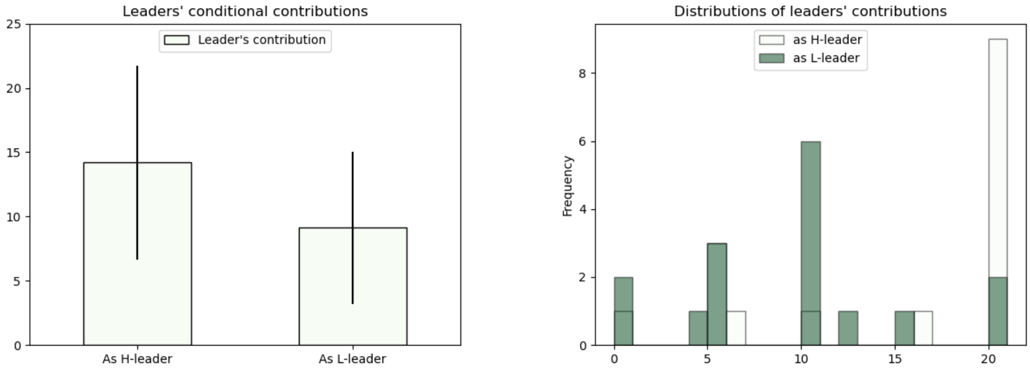


(a) With respect to H-leaders (b) With respect to L-leaders

Figure II.2: Conditional contributions of followers (endogenous returns)

It seems that the differences in realized contributions are not driven by differences in followers' contribution preferences. Therefore, they must stem from different contributions of the leader. Figure II.3 below shows leaders' contributions in the one-shot game and their distributions, depending on the leader's return. As we see, leaders are willing to contribute more if they have a high return.

Result 13. *Leaders are willing to contribute more if the followers assign them a high return from the public good.*



(a) Mean conditional contributions of the leaders (with st. devs.)

(b) Distributions of leaders' conditional contributions

Figure II.3: Leaders' conditional contributions and their distributions (endogenous returns)

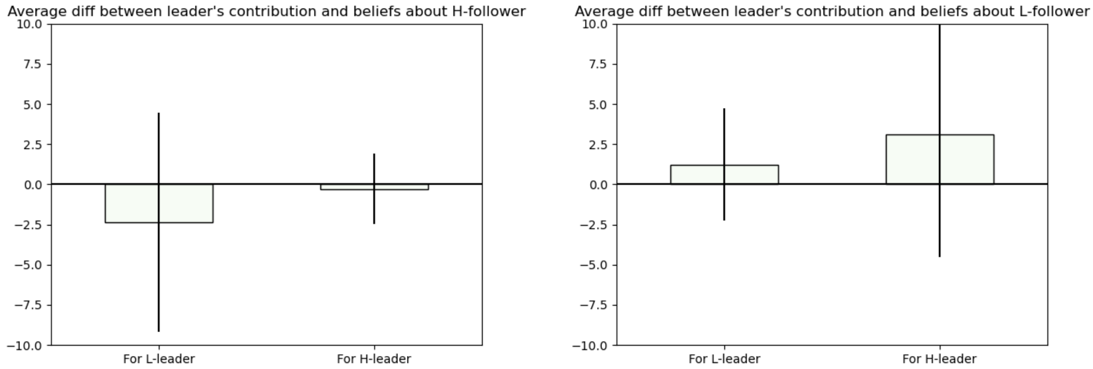
Why would leaders contribute more in case they receive more from the public good? There are two reasons for that. First, leaders might reciprocate the voting decisions, and increase their own contributions to appreciate the voting decisions of the followers.[14] Second, it might be that leaders anticipate different levels of cooperation from followers in cases when the leader has low or high returns from the public good. Leaders might realize that it is worth cooperating with the H-follower when the leader has a high return from the public good: if both H-players contribute their full endowment, they receive more than they do if they contribute nothing, regardless of the contribution by L-follower. Figure II.4 suggests that leaders expect followers with different returns to contribute differently, but no differences in leaders' expectations about followers' contributions are observed regarding the leader's own return (Wilcoxon signed-rank tests $p > 0.27$).

Result 14. *Leaders expect that H- and L-followers contribute differently (in response to identical levels of leaders' contributions).*

We have already seen that leaders' expectations coincide with the observed differences in the contributions of followers. Figure II.5 below shows that followers hold quite similar beliefs about the contribution of the other follower. However, H-followers expect L-followers to contribute on average more than L-followers expect H-followers to do in response to low (below 10) leaders' contributions.

These expectations seem surprising. Figures II.6 and II.7 demonstrate how followers' conditional contributions are related to their own beliefs about contribution of the other follower (conditional on the leader's contribution and the leader's return from the public good). High expectations of H-followers about contributions of L-followers in groups with H-leaders could be attributed to their preference for equal contributions rule (panel II.6a); nevertheless, there is no such justification in groups with L-leaders (panel II.7a). More-

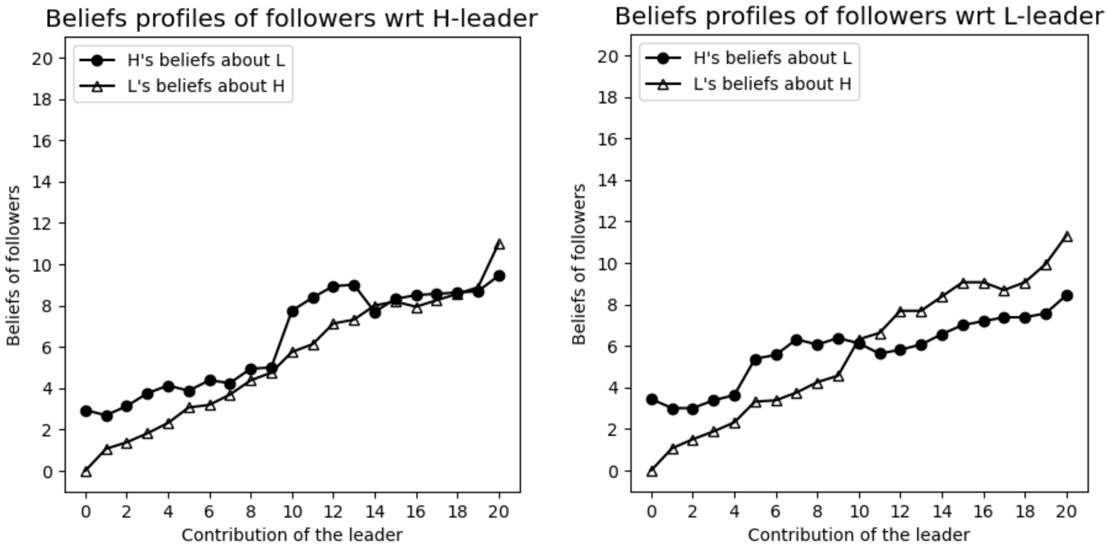
[14] Furthermore, their contributions might be driven by guilt aversion.



(a) About H-followers (b) About L-followers

Figure II.4: Leader's contribution net leader's respective belief (endogenous returns)

Note: The bars are to be interpreted as follows. Positive numbers reflect that the leader contributes more than she expects a follower with a particular return to contribute, while negative number reflects that the leader contributes less than the expected contribution of a follower. Vertical lines correspond with standard deviations.

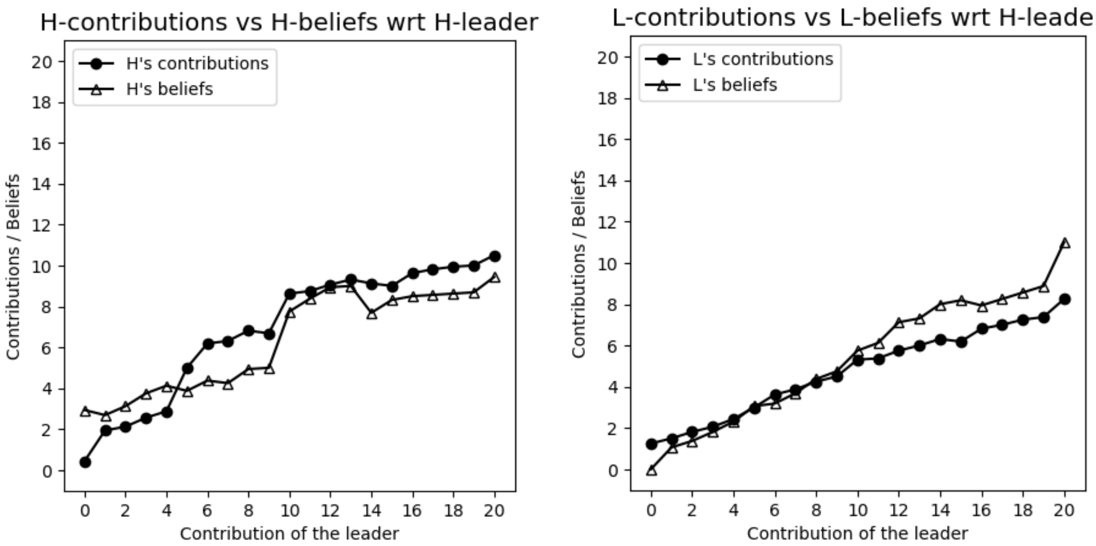


(a) W.r.t. H-leader's contribution (b) W.r.t. L-leader's contribution

Figure II.5: Averaged beliefs profiles of followers with respect to identical leaders (endogenous returns)

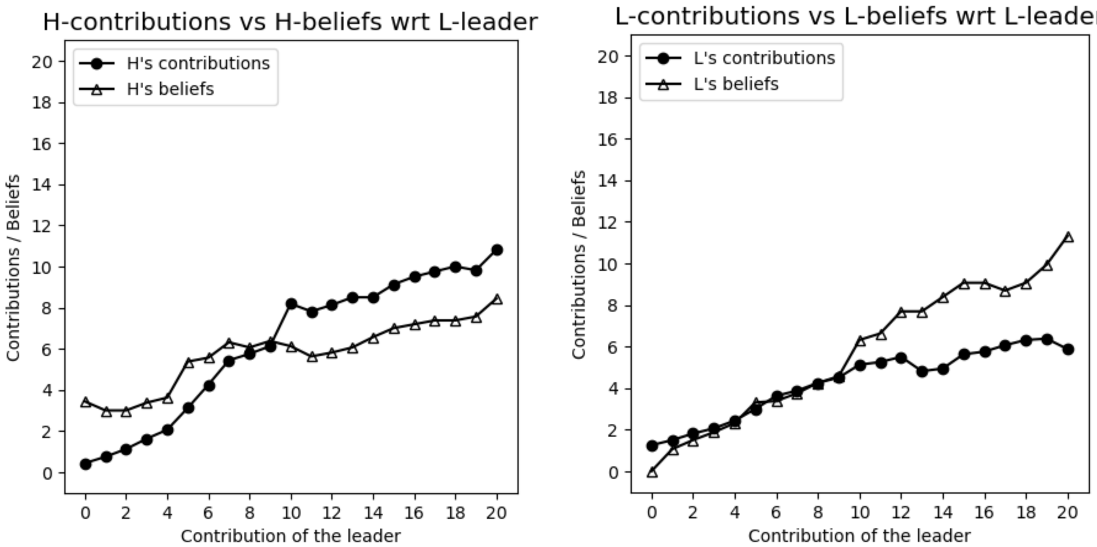
over, H-followers expect positive contributions from L-followers even for negligible levels of leaders' contributions.

Result 15. *Followers' contributions almost coincide with their own beliefs about the contribution of the other follower in response to the contribution of potential H-leaders but not in case of L-leaders. In the latter case, when leaders' contributions are large, L-followers contribute substantially less than they expect H-followers to contribute, and H-followers contribute more than they expect L-followers to contribute.*



(a) H-followers (b) L-followers

Figure II.6: Averaged contribution profiles and own beliefs about the other follower with respect to the *H-leader*'s contribution decision (endogenous returns)



(a) H-followers (b) L-followers

Figure II.7: Averaged contribution profiles and beliefs with respect to the *L-leader*'s contribution decision (endogenous returns)

II.5.1 Regression Analysis

In order to more rigorously analyze the differences in contribution patterns of followers with respect to potential leaders with different returns, I use the following econometric model:

$$\begin{aligned}
C_{i,j} = \alpha_0 + \beta_0 \cdot C_l(j) + \gamma_0\ B_{i,j} \\
&+ \alpha_{leader}\ H\text{-}leader + \alpha_{own}\ L\text{-}follower + \alpha_{HL}\ H\text{-}leader \times L\text{-}follower \\
&+ \beta_1\ H\text{-}leader \times C_l(j) + \beta_2\ L\text{-}follower \times C_l(j) \\
&+ \beta_3\ H\text{-}leader \times L\text{-}follower \times C_l(j) \\
&+ \gamma_1\ H\text{-}leader \times B_{i,j} + \gamma_2\ L\text{-}follower \times B_{i,j} + \gamma_3\ H\text{-}leader \times L\text{-}follower \times B_{i,j} \\
&+ \alpha_X\ X_i + \varepsilon_{i,j}, \qquad \text{(II.1)}
\end{aligned}$$

where $C_{i,j}$ is a dependent variable reflecting the conditional contribution in response to the leader's contribution, $C_l(j) \equiv j$; $B_{i,j}$ is the own belief about the contributions of the other follower in response to the leader's contribution, $L\text{-}follower$ is a dummy equal to 1 if the follower has a low return from the public good and 0 otherwise, $H\text{-}leader$ is a dummy equal to 1 if the potential leader has a high return (i.e. 1 for H-HL groups) and 0 otherwise (for L-HL groups), X_i is the set of individual characteristics from the questionnaire, and $\varepsilon_{i,j}$ is a random effects error term.

Table II.3 presents the pooled sample estimates for four econometric specifications of random effects models for conditional contributions of the followers. Column 1 presents the simplest specification, which is a RE model for own (conditional) contribution with respect to the leader's contribution, dummies reflecting the composition of returns in the group, and the set of interaction terms to capture systematic differences in patterns. The model presented in column 2 adds the follower's own beliefs of the followers about (conditional) contributions of the other follower in the group as an additional variable, along with corresponding interaction terms capturing heterogeneity in beliefs. The model presented in column 3 also includes a dummy variable equal to 1 if the follower's voting decision was for the leader to have a high return, and also interacted with a dummy on the follower's low return. The model presented in column 4 additionally includes individual characteristics from the questionnaire, namely gender, self-assessed risk preferences, trust, and power distance.[15] All standard errors are clustered at the individual (participant's) level.

As Table II.3 demonstrates, the leader's contribution is a statistically significant explanatory factor for the contribution of the follower, as well as the followers' own beliefs about contribution of the other follower in the group, but there are also substantial differences depending on the composition of returns in the groups. It seems that L-followers do not respond to H-leaders' contributions at all.

In order to ascertain whether followers follow leaders with varying returns differently I estimate the same regression separately for the sub-samples of H- and L-followers (see Table II.B.1 in Appendix II.B). The results presented in Table II.B.1 show substantial

[15] Individual characteristics were transformed into dummies reflecting gender, high risk aversion, high risk loving, very high/low trust, and strong/weak power distance. These additional variables are not reported in the regressions below.

Table II.3: Estimates for random effects models of conditional contributions (endogenous returns)

	(1) Contribution	(2) Contribution	(3) Contribution	(4) Contribution
Leader's contribution	0.486**	0.297*	0.297*	0.297*
	(0.155)	(0.118)	(0.118)	(0.118)
L-follower	-0.701	0.462	3.536	4.161
	(1.744)	(1.315)	(2.612)	(4.397)
H-leader	-1.261	-0.714	-0.715	-0.715
	(1.090)	(0.703)	(0.703)	(0.704)
L-follower × Leader's contribution	-0.143	-0.167	-0.167	-0.165
	(0.185)	(0.149)	(0.149)	(0.150)
L-follower × H-leader	1.637	1.163	1.164	1.164
	(1.128)	(0.775)	(0.775)	(0.776)
H-leader × Leader's contribution	0.060	0.191	0.191	0.191
	(0.120)	(0.132)	(0.133)	(0.133)
L-follower × H-leader × Leader's contribution	-0.145	-0.239	-0.239	-0.239
	(0.135)	(0.135)	(0.136)	(0.136)
Belief		0.513***	0.513***	0.513***
		(0.110)	(0.110)	(0.110)
H-leader × Belief		-0.273*	-0.272*	-0.272*
		(0.119)	(0.119)	(0.119)
L-follower × Belief		-0.080	-0.080	-0.084
		(0.196)	(0.196)	(0.198)
L-follower × H-leader × Belief		0.169	0.168	0.168
		(0.145)	(0.145)	(0.146)
VotedHigh			2.427	4.602
			(2.486)	(3.060)
VotedHigh × L-follower			-3.859	-3.317
			(2.996)	(4.255)
Constant	2.029	0.674	-1.148	-3.421
	(1.177)	(0.682)	(2.276)	(3.690)
Other controls	No	No	No	Yes
Observations	1344	1344	1344	1344
Number of individuals	32	32	32	32

Notes: *L-follower* is a dummy equal to 1 if the follower has a low return from the public good; *H-leader* is a dummy equal to 1 if the leader in the group was selected to have a high return from the public good; *Belief* is the belief from a conditional beliefs table about how much the other follower in the group contributes in response to a particular contribution by the leader; *VotedHigh* is a dummy equal to 1 if the follower voted for the leader to have a high return. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

differences in contribution patterns of the followers with different returns. For example, the leader's contribution is the only important factor for H-followers: once beliefs about the contribution of the other follower in the group are included (columns 2 and 6). It also appears that H-followers follow leaders more closely than L-followers, holding the rest equal. Voting for the leader to have a high return from the public good (columns 3 and 7) increases contributions of H-followers but not L-followers. Adding individual controls (columns 4 and 8) does not change the results.[16] Large negative coefficients for the intercept in columns 4 and 8 suggest that there is most likely an endogeneity problem between

[16] Tables II.B.2, and II.B.3 in Appendix II.B also present estimates for a fixed effects model of conditional contributions, similar to (II.1). These tables show the robustness of the results reported above.

voting decision and individual characteristics.[17] Overall, beliefs about the other follower's contributions seem to be the most important driver for followers' contribution decisions.

Result 16. *Although most of followers voted for the leader to have a high return from the public good, the effect of leaders' contributions on the contribution decisions of L-followers is not statistically significant.*

Result 17. *Beliefs about the other follower's contribution seem to be a key driver for followers' own contributions.*

The results presented in Tables II.3 and II.B.1 do not allow for a clear comparison in reactions to "alike" leaders. To make this comparison more explicit, I estimate the Random effects model for follower's conditional contributions with respect to the contribution of the leader who has the same or different returns from the public good. The results are reported in Table II.4 below.

Table II.4: Estimates for random effects model of conditional contributions with respect to contributions by the leader with similar/different return (endogenous returns)

	(1) H-followers	(2) L-followers
Leader with same return	-0.716	-0.448
	(0.719)	(0.333)
Contribution by leader with same return	0.488***	0.134
	(0.143)	(0.094)
Contribution by leader with different return	0.297*	0.086
	(0.120)	(0.097)
Belief, when leader has same return	0.241	0.425*
	(0.139)	(0.169)
Belief, when leader has different return	0.513***	0.322*
	(0.112)	(0.145)
VotedHigh	4.748	3.076
	(3.511)	(3.790)
Constant	1.467	-2.227
	(5.316)	(5.293)
Other controls	Yes	Yes
Observations	672	672
Number of individuals	16	16

Notes: *Belief* is the belief from a conditional beliefs table about how much the other follower in the group contributes in response to a particular contribution by the leader; *VotedHigh* is a dummy equal to 1 if the follower voted for the leader to have a high return. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

Table II.4 demonstrates an important finding. L-followers do not seem to rely on the leader's contribution at all, but base their contribution decision on their own belief

[17] See also footnote 19 below.

about the contribution of the other follower in the group.[18] Furthermore, voting for the leader to have a high return on average substantially reduces the contributions of L-followers but increases the contributions of H-followers. Interestingly, H-followers respond to the contribution of H-leaders more than for L-leaders (unlike L-followers, who do not respond to the leader's contribution at all), providing evidence for their willingness to profit from cooperation (intercept dropped in column 2 due to collinearity issues).[19]

Result 18. *H-followers rely on both the contribution of the leader and their own belief, while they also respond similarly to the contribution of L- and H-leaders. At the same time, L-followers do not seem to respond to the contribution of the leader at all, relying on their own beliefs about the contribution of the other follower in the group.*

Nevertheless, this effect is strongly associated with the voting for the leader to have a high return (see Table II.5 below): L-followers who do *not* vote for H-leaders respond quite generously to both H- and L-leaders' contributions. At the same time, there is no such relation between voting decisions of H-followers and their contributions.

Result 19. *Voting decisions shape L-followers' conditional contributions: L-followers who vote for the leader to have a low return follow both H- and L-leaders similarly; furthermore, their responses to the leader's contributions are similar to those of H-followers. There is no such relation between voting decisions of H-followers and their contributions.*

However, it is unclear whether voting decisions drive contributions or contribution preferences determine voting decisions.[20]

II.6 Repeated Game

II.6.1 Motivation

When the game is played repeatedly, players' decisions might change due to strategic considerations, raising new questions about the effectiveness of the institution. Do followers' voting decisions in a repeated game differ from those in a one-shot game due to the possibility for strategic cooperation? Do they respond to actual leaders' contributions in a similar way as they did in the one-shot game? In order to answer these questions I check how contribution behavior in a one-shot game translates into dynamics in a repeated interaction.

[18] Table II.B.4 in Appendix II.B confirms the robustness of the result: L-followers do not rely on the leader's contribution, but rather their own belief about the contribution of the other follower in the group.

[19] Gender, trust and risk preferences predict voting decisions among L-followers: being a low-trusting female who does not assess herself as a low risk-averse predicts L-followers to vote for an H-leader. For H-followers, power distance plays a role.

[20] In the experiment, players decided on contributions without knowing the outcome of voting, although the voting stage preceded the contributions stage. In the contributions stage, players filled out tables conditional on the *voting outcome*, not on their own voting decision.

Table II.5: Estimates for random effects model of conditional contributions with respect to voting decisions (endogenous returns)

	(1) H-followers	(2) L-followers
Leader with same return	-0.711	-0.452
	(0.680)	(0.332)
Contribution by leader with same return	0.210	0.493***
	(0.231)	(0.064)
Contribution by leader with different return	0.105	0.452**
	(0.196)	(0.142)
Belief, when leader has same return	0.176	0.441**
	(0.152)	(0.164)
Belief, when leader has different return	0.432***	0.329*
	(0.118)	(0.140)
VotedHigh	1.091	7.234*
	(3.311)	(3.548)
VotedHigh × Leader's Contribution (same return)	0.391	-0.420***
	(0.244)	(0.084)
VotedHigh × Leader's Contribution (different return)	0.296	-0.423**
	(0.251)	(0.146)
Constant	4.912	-5.885
	(5.540)	(4.851)
Other controls	Yes	Yes
Observations	672	672
Number of individuals	16	16

Notes: $Belief$ is the belief from a conditional beliefs table about how much the other follower in the group contributes in response to a particular contribution by the leader; $VotedHigh$ is a dummy equal to 1 if the follower voted for the leader to have a high return. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

In this section, I thus report the results from the second part of the experiment, where the same subjects who played the one-shot game were asked to play in the same groups as before for ten consecutive rounds of the public goods game. I find that the majority of players decide to vote for the leader to have a high return. Leaders that were elected to have a low return started with full contributions, but low responses from followers led to an overall poor cooperation in L-HL groups. In general, the patterns observed in the repeated game are similar to those observed in the one-shot game, namely that L-followers are less willing to cooperate. The regression analysis suggests that both H- and L-followers follow their leaders much more closely in the repeated game than in the one-shot game.

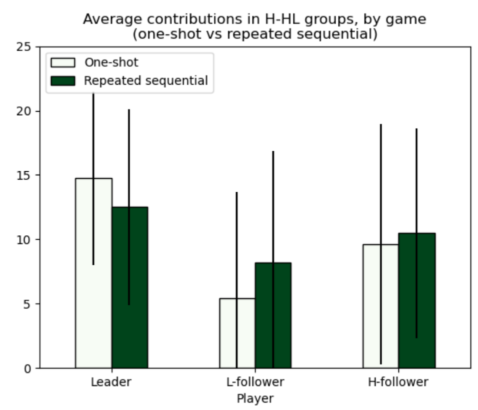
II.6.2 Results for the Repeated Games

I start with presenting realized outcomes for the repeated game: voting decisions of followers and realized contributions and payoffs. Table II.6 below shows summary statistics on voting decisions of the followers in the repeated game. Again, more than 80% of all followers voted for their leader to have high return from the public good, and there is a similar heterogeneity in voting decisions between H- and L-followers as in the one-shot game. There are only three subjects who voted "low" in the one-shot game that also voted "low" in the repeated game.

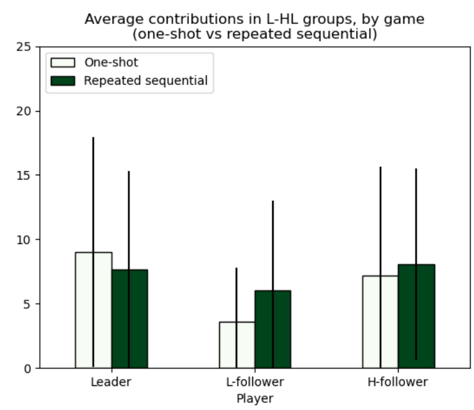
Table II.6: Voting decisions of followers in the repeated game, %

	Voted	
Follower	**Low**	**High**
L-follower	12.5	87.5
H-follower	25	75

Table II.C.1 in Appendix II.C shows basic summary statistics (mean, standard deviation, and median) on actual contributions and payoffs in repeated games, separately for periods 1, 2–5, 6–9, 10, and over the whole game. Only three groups had leaders with low returns in the repeated game (compared with thirteen groups who had leaders with high returns). It is clear that all players contribute more in the first round of the repeated game than what they stated in their conditional contributions. However, as the game proceeds we can observe different patterns: in L-HL groups, contributions rapidly decline, and already in rounds 2–5 they drop below the levels reported in the one-shot game; however, for H-HL groups players were able to maintain much higher levels of cooperation, as average contributions do not fall below the levels reported in the one-shot game until the last round. In both one-shot and the repeated games, average contributions are lower in groups with an L-leader, as Figure II.8 demonstrates.



(a) H-HL groups



(b) L-HL groups

Figure II.8: Differences in average contributions between one-shot and repeated games (endogenous returns)

Result 20. *Groups that elected H-leaders contribute more than groups that elected L-leaders and sustain higher levels of cooperation during the course of the game.*

Because in the repeated game voting occurs only once at the beginning of the game and it precedes contributions, it might seem that voting decisions cause contributions in repeated interaction. However, the regression analysis below suggests the opposite (see the next subsection).

Leaders increase their initial contributions in the first round of the repeated game, as Figure II.9 demonstrates. This effect is most salient for L-leaders.

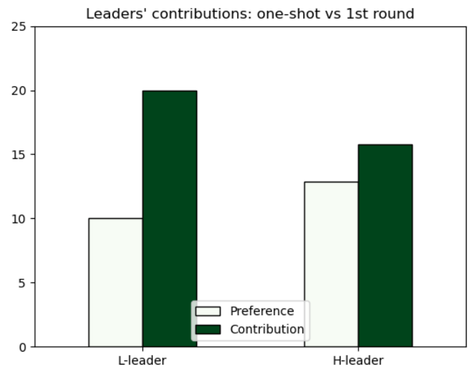
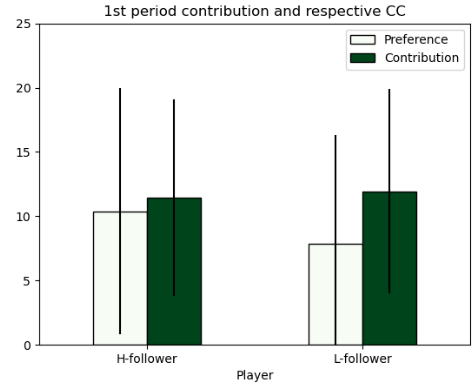


Figure II.9: Leaders' conditional contribution (preference) from the one-shot game and corresponding contribution in the first period of the repeated game (endogenous returns)

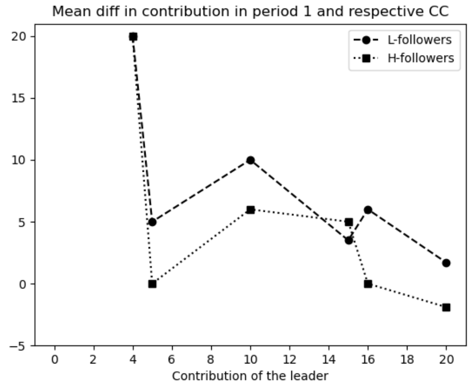
Do followers increase their contributions strategically in the repeated game compared with what they stated in their contribution preference in the one-shot game? Figure **??** shows that they do. Graph on the left (panel II.10) shows average actual contributions of the followers in the first period of the repeated sequential game and their respective contribution extracted from the conditional contribution tables, given the actual contribution of the leader in the first period of a repeated game. The differences are not statistically significant (Wilcoxon signed-rank test $p > 0.24$). The graph on the right (panel II.10b) shows average deviations from the stated conditional contribution that corresponds to actual leaders' contribution in the first round. It appears that followers do increase their contributions compared with their contribution preferences stated in the one-shot game, at least for low levels of leaders' contributions.

Result 21. *Players seem to take strategic considerations into account: all players increase their contribution decisions in the first period of the repeated sequential game, compared with what they stated in the contribution tables in the one-shot game.*

A more detailed look at contribution dynamics in different groups (see Figure II.11) suggests that poor cooperation in L-HL groups is likely to be driven by low contributions of followers: in the first period of the repeated game, followers in L-HL groups contributed almost the same amount as followers in H-HL groups, despite all leaders contributed fully in L-HL groups. This could be viewed by the leaders as non-reciprocation of their high contribution ("sacrifice") and resulted in rapid decrease in contribution levels in later periods.

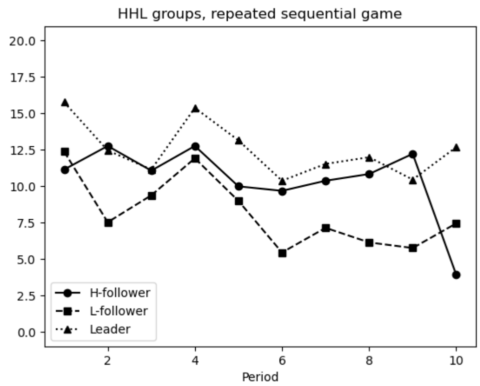


(a) Average contributions

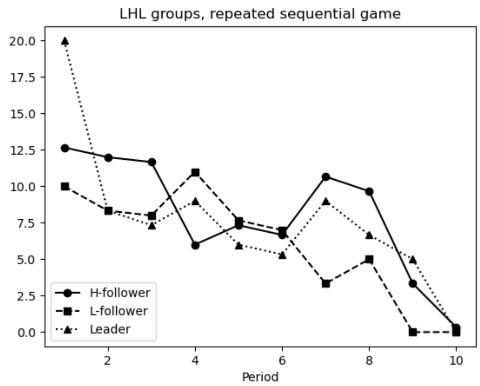


(b) Average differences

Figure II.10: Comparison of average contribution in the first period of repeated sequential game and corresponding contribution from the schedule (endogenous returns)



(a) H-HL groups



(b) L-HL groups

Figure II.11: Dynamics of contributions (endogenous returns)

The leaders in L-HL groups respond immediately to low followers' contributions in the second round and lower their contributions as well, but only after they observed followers' reactions in the first round. Therefore, low cooperation cannot be attributed to leaders' "disappointment" from having a low return.

Result 22. *Low cooperation among groups that selected the low leader's return is driven by low contributions of the followers and not caused by the leader's "disappointment" about receiving a low return.*

II.6.3 Regression Analysis

Below I analyze what drives players' contributions in repeated games. Since I do not elicit players' beliefs about contributions of others in repeated games, I use the values imputed from the conditional beliefs schedule from the one-shot game with the corresponding leader's contribution to proxy players beliefs.

I estimate the following dynamic panel regression for the contributions of followers in the repeated sequential game:

$$\begin{aligned} C_{i,t} = {} & \alpha_0 + \alpha_{leader} L_t + \alpha_{belief} B_{i,t} \\ & + \alpha_{leader}\ H\text{-}leader + \alpha_{own}\ L\text{-}follower + \alpha_{HL}\ H\text{-}leader \times L\text{-}follower \\ & + \beta_1\ H\text{-}leader \times L_t + \beta_2\ L\text{-}follower \times L_t + \beta_3\ H\text{-}leader \times L\text{-}follower \times L_t \\ & + \delta_{leader} \Delta P_{leader,t-1} + \delta_{other} \Delta P_{-i,t-1} + \alpha_X X_i + \alpha t + \varepsilon_{i,t}, \end{aligned} \tag{II.1}$$

where $C_{i,t}$ is dependent variable reflecting the own contribution in the current period t, L_t is the leader's contribution in the current period, $B_{i,t}$ is the belief imputed *from the conditional beliefs schedule from the one-shot game* that corresponds to the leader's contribution in period t, $H\text{-}leader$ and $L\text{-}follower$ are dummies reflecting the composition of returns in the group, $\Delta P_{leader,t-1}$ is the difference between the own payoff and the leader's payoff in the previous period, $\Delta P_{-i,t-1}$ the difference between the own payoff and the payoff of another follower in the group in the previous period, X_i is the set of individual characteristics, and $\varepsilon_{i,t}$ is a random effects error term.

For leaders' contributions, I estimate the following equation:

$$L_{j,t} = \omega_0 + \omega_1 H\text{-}leader + \omega_H \Delta P_{H,t-1} + \omega_L \Delta P_{L,t-1} + \omega t + \omega_X X_j + \eta_{j,t}, \tag{II.2}$$

where $L_{j,t}$ is the dependent variable reflecting the leader's contributions in the current period, t, *H-leader* is a dummy reflecting the leader having a high return from the public good, $\Delta P_{H,t-1}$ is the difference between the own payoff and the payoff of H-follower in the previous period, $\Delta P_{L,t-1}$ is the difference between the own payoff and the payoff of L-follower in the previous period, X_j is the set of individual characteristics, and $\eta_{i,t}$ is a random effects error term. Equations (II.1) and (II.2) were estimated separately.

Table II.7 below shows estimates for a dynamic panel random effects model for followers' contributions in the repeated game. All standard errors are clustered at the group level (of three participants).

Table II.7 demonstrates that in the repeated game, after additional individual controls are included, in their contribution decision followers take into account the contribution of the leader and their expectation about the contribution of another follower in their group (proxied by a corresponding belief from the belief schedule elicited in the one-shot game). The difference between the own payoff in the previous period and the corresponding leader's payoff affects the followers' contributions, whereby the more advantageous that the difference is, the lower the follower's contribution. L-followers contribute on average much less than H-followers. The effect of the leader's contribution varies for followers with different returns, whereby L-followers react less to the contribution of the leader than H-followers. Moreover, the effect of the leader's contribution varies depending on the leaders' return: on average, followers follow H-leaders less closely. Followers' contributions decline over time. Interestingly, own voting decisions do not have any statistically significant

Table II.7: Estimates for random effects model for dynamic contributions of followers (endogenous returns)

	(1) Contribution	(2) Contribution
Leader's contribution(t)	0.991***	0.922***
	(0.093)	(0.121)
Period, t	-0.405***	-0.411***
	(0.115)	(0.115)
Belief from one-shot game	0.199	0.184
	(0.113)	(0.127)
Difference with leader's profit($t-1$)	-0.368***	-0.311***
	(0.085)	(0.086)
Difference with the profit of the other follower($t-1$)	0.086	0.078
	(0.061)	(0.062)
L-follower	-0.955	-3.221
	(0.592)	(1.875)
H-leader	-1.929**	-3.969**
	(0.669)	(1.390)
L-follower × Leader's contribution(t)	-0.537***	-0.436***
	(0.072)	(0.117)
H-leader × Leader's contribution(t)	-0.404***	-0.365*
	(0.116)	(0.172)
L-follower × H-leader	-0.524	1.393
	(1.796)	(2.110)
L-follower × H-leader × Leader's contribution(t)	0.361*	0.317
	(0.175)	(0.203)
VotedHigh	1.531*	2.647
	(0.643)	(1.443)
VotedHigh × L-follower	0.272	-0.918
	(1.532)	(2.175)
Constant	4.801***	6.888***
	(0.825)	(1.287)
Other controls	No	Yes
Observations	288	288
Number of individuals	32	32

Notes: *L-follower* is a dummy equal to 1 if the follower has a low return from the public good; *H-leader* is a dummy equal to 1 if the leader in the group was selected to have a high return from the public good; $Difference\ with\ leader's\ profit(t-1)$ and $Difference\ with\ profit\ of\ another\ follower(t-1)$ reflect differences in profits of the follower and respective players in the previous period. $VotedHigh$ is a dummy equal to 1 if the follower voted for the leader to have a high return in the repeated game. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

effect once individual controls are included. The latter observation speaks in favor of the hypothesis that voting decisions represent the signaling for the leader with a preferred return (to allow followers to contribute according to their preferences), and do not cause contribution decisions.

Result 23. *Contributions of followers in a repeated sequential game are almost exclusively explained by the leader's contribution, their own beliefs about contribution of another follower in the group, and the composition of returns in the group.*

Result 24. *Contributions of followers in a repeated game are less generous in response to the contributions of H-leaders. Similarly to the one-shot game, L-followers respond less generously to the leader's contributions, regardless of the leader's return.*

Result 25. *Contributions of followers decline over time.*

Table II.8 shows estimates for a dynamic panel random effects model for leaders' contributions in a repeated game. All standard errors are clustered at the group level (of three participants).

Table II.8: Estimates for random effects model for dynamic contributions of leaders (endogenous returns)

	(1) Contribution	(2) Contribution
Difference with profit of L-follower($t-1$)	0.166*	0.167*
	(0.069)	(0.069)
Difference with profit of H-follower($t-1$)	0.106	0.100
	(0.118)	(0.123)
H-leader	4.435	5.809
	(3.690)	(3.122)
Period, t	-0.238	-0.234
	(0.223)	(0.228)
Constant	8.627*	-2.982
	(3.872)	(3.601)
Other controls	No	Yes
Observations	144	144
Number of individuals	16	16

Notes: *H-leader* is a dummy equal to 1 if the leader in the group was selected to have a high return from the public good; *Difference with profit of L-follower*($t-1$) and *Difference with profit of H-follower*($t-1$) reflect differences in profits of the leader and respective followers in the previous period. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

According to Table II.8, on average leaders take only the difference between their own payoff in the previous period and the corresponding payoff of L-follower into account. On average, leaders' contributions do not fall over time, and H-leaders contribute substantially more than L-leaders, given the rest equal.

Result 26. *Leaders' contributions are only affected by the difference in payoffs between the leader and L-follower in the previous period but do not fall over time per se. H-leaders contribute substantially more than L-leaders.*

The results obtained in the repeated game are in line with the findings in the literature on this topic. First, with repeated interaction strong path dependence occurs. Leaders' contributions play a major role in determining these contribution paths, as followers are very responsive to leaders' contributions in repeated games. Second, overall contributions decline over time, which occurs, according to Fischbacher and Gächter (2010) due to the fact that players are imperfect conditional cooperators. Interestingly, I find that leaders' contributions do not decline per se but react to reduced followers' contributions, which may suggest that leaders try to act as a role model during the whole length of the experiment (see Gächter and Renner, 2018). Third, all players (both leaders and followers) in repeated sequential games rely on differences in payoffs among players in the previous period (which suggests the presence of some sort of social comparison); however, reference groups seem to differ. While followers take into account the difference between their own payoff and the payoff of the leader, leaders compare their payoffs with the payoff of L-followers (the most "disadvantaged" players). Finally, heterogeneity in returns affects contributions in the repeated sequential game in a similar way that we observed it in the one-shot game, although, followers seem to be more responsive to leaders' contributions in the repeated interaction. This responsiveness might occur due to more precise updating in the repeated game or due to better coordination on a particular contribution norm; however, I am unable to distinguish between these two hypotheses.

II.7 Concluding Discussion

Leadership is typically considered to be a mechanism that calls leaders for potential sacrifices (e.g. putting at risk their own self-interests). A central motive for leaders to contribute in a sequential public goods game, if there are no strategic considerations, is their expectations about followers' reciprocity. The literature has established that with heterogeneity in returns, players seem to follow different contribution rules. The sequential nature of the game therefore implies that different contribution norms might be established in groups led by leaders with varying returns (i.e. with diverse compositions of returns in the group). However, there is limited understanding of the functioning of leading-by-example if the composition of returns is endogenous, i.e. when it is determined by (a subgroup of) the players.

In this paper I studied in an economic experiment how reciprocal reactions of the followers (second-movers in a sequential public goods game) and leaders' behavior change when followers determine the composition of returns by voting on the return that the leader gets. This can be viewed as a possibility to choose between different group compensation schemes in collaborative projects.[21] For example, would workers exert more or less effort if they are able to choose a technology that provides additional benefits for other mem-

[21] One can also view this as an attempt to mimic the selection of a group between two types of groups to be a member of: one type has the first-mover (the leader) with a high return from the public good, and another type has the leader with lower return.

bers, like rewarding project managers with additional compensation? The within-subject experimental design allowed me to analyze reactions to changing the leader's return from the public good using conditional contributions in a one-shot game. Additional results from the repeated interaction allowed me to investigate how these contribution decisions translate into dynamics.

My findings suggest that the voting decisions of the followers line up with the hypothesis that followers expect leaders with high returns to contribute more. Leaders indeed contribute larger amounts when they receive high returns in the one-shot game. Nonetheless, when it comes to a repeated interaction, this relation is reversed in the first round of the repeated game. In the repeated game, followers decide once on the leader's return from the public good, and the resulting composition of returns is kept throughout the game. Leaders who have been selected to have a low return did not demonstrate any disappointment about receiving a low return and started with full contributions, but followers in such groups did not reciprocate leaders' contributions. As a result, groups that ended up with leaders with low returns demonstrate on average lower provision of the public good and a rapid decrease in cooperation over time.

Although more than 80% of the followers vote for the leader to have a high return, their reactions to leaders' contributions differ, depending on their own return from the public good. While followers with high returns follow the actual leader's contribution quite closely in the one-shot game, followers with low returns start free-riding more. It seems that followers try to motivate their leaders to contribute larger amounts, appreciating receiving high returns. Regression analyses suggest that differences in followers' responses are closely related to their voting decisions. For high-benefiting followers (H-followers), their votes do not change their contribution responses, but responses of low-benefiting followers (L-followers) change dramatically. Those L-followers who vote for the leader to have a low return contribute quite generously regardless of the resulting voting outcome.

These changes may be related to the justification of asymmetries between group members. Van Dijk and Wilke (1993) show that, when the asymmetry in the interest from (share of) the public good is not justified, the high-interest subjects contribute more than the low-interest subjects, but when the asymmetry is justified, no differences in contributions were observed. Later, Cherry et al. (2005) studied the justification effect, varying the origin of endowments in a linear public goods game (as do Kroll et al., 2007, in a best-shot public goods game): earned versus windfall. They find that people assign different weights to fairness considerations: an efficient outcome might not seem fair in asymmetric groups with earned endowments. Manipulating the justification of group asymmetries may change the results on the effect of leading-by-example in my setup.

The observed patterns of contributions suggest that followers behave as if they avoid being the "sucker"; i.e. followers aim to ensure they do not receive the lowest payoff in the group.[22] However, this assumption does not explain the absence of L-followers'

[22] In this case, H-followers can be certain they do not receive the lowest payoff if they do not exceed leaders' contributions. However, for L-followers to ensure they do not receive the lowest payoff, they should

reactions to contributions made by potential L-leaders when these followers voted for the leader to have a high return. A plausible explanation seems to be that such followers try to avoid being the "sucker" in the group and maximize their profits at the same time.

My findings underline the importance of belief management in policies aimed to increase the provision of public goods: it seems that, if some group members consider themselves as disadvantaged ones, they base their decisions solely on their expectations and (perceived) differences but not on the observed actions of other members, and thus they tend to free-ride more. They also suggest that introducing democratic elections may have an ambiguous effect on the provision of public goods: on the one hand, this might backfire the intended "role model" effect (if some group members decide to perceive themselves as "disadvantageous"), but on the other hand, it might help to motivate other participants to engage in the public goods provision.

contribute no more than a half of the contribution of H-leaders.

APPENDICES

II.A Experimental Instructions

These instructions were translated from German.

II.A.1 General Instructions for the Participants

You are now taking part in an economic study. If you read the explanations below carefully then you can — depending on your decisions and those of other participants — earn a not inconsiderable amount of money. It is therefore very important that you carefully read and understand these explanations.

During the study there is an absolute ban on communication. If you have any questions, please address them to us only: raise your hand and we will come to you. Failure to comply with this rule will result in exclusion from the trial and all payments.

The study comprises of several independent parts. In the following the course of the first part will be described first. Other parts will be described later. For today's participation, you will receive a show-up fee of 4 Euros. Your income is increased in each part of the study. However, during the study we are not talking about Euros, but about Talers. Your total income is thus initially calculated in Talers. The total number of Talers you earn during the study will then be converted into Euros and paid out in cash. The exchange rate from Taler to Euro will be announced at the beginning of the respective part.

II.A.2 Instructions for Part 1 (paper)

II.A.2.1 General Information

In the first part of the study, the exchange rate is: 10 Talers = 6 Euro.

The first part of the study comprises of a **single period**. At the beginning of the first part, you are randomly divided into groups of three. In each group, there are participants of type B and of type A.

Each participant receives 20 Talers. Your task now is to decide on the use of your 20 Talers. You can contribute all or part of the 20 Talers to a **project** or put it on a **private account**. Any Taler that you do not contribute to the project will be automatically added to your personal account. For example, if your contribution to the project is 5 Talers, 15 Talers will remain in your private account.

Income from the private account:

For every Taler that you put on the private account, you earn exactly one Taler. For example, if you place 20 Talers on the private account (and thus contribute nothing to the project) you will earn exactly 20 Talers from the private account. If you contribute, for example, 12 Talers to the project (putting 8 Talers in your personal account), you will receive an income of 8 Talers from the private account. *No one except you receives any income from your private account.*

Income from the project:

For every Taler that you or another participant in your group contributes to the project, *you and all other participants in your group* receive a certain income from the project. How much income each participant receives from the project depends on which type he or she is.

Type A: income from the project = sum of contributions to the project × 0.5
Type B: income from the project = sum of contributions to the project × 0.75

Examples: If the sum of the contributions of all participants to the project is 15 Talers (e.g. if you and the other two participants each contribute 5 Talers), then a participant of type B in your group will receive $15 \times 0.75 = 11.25$ Talers from the project, and a participant of type A in your group receives $15 \times 0.5 = 7.5$ Talers. If instead the sum of the contributions to the project is 10 Talers in total, a participant of type B in your group will receive $10 \times 0.75 = 7.5$ Talers from the project, and a participant of type A in your group will receive $10 \times 0.5 = 5$ Talers, etc.

Your pay-out from the first part is calculated by adding up your income from your personal account and your income from the project:

Type A:
Income from the private account (= 20 – contribution to the project)
+ Income from the project (= 0.5 × sum of contributions to the project)
Payment from the first part of the study

Type B:
Income from the private account (= 20 – contribution to the project)
+ Income from the project (= 0.75 × sum of contributions to the project)
Payment from the first part of the study

Procedure:

The first part of the study comprises of several stages, which are explained below.

Stage 1: At the first stage, one participant is randomly assigned type A, and another participant is assigned type B.

Stage 2: These two participants then decide whether the third participant should be of type A or type B. Each participant can either cast a vote for type A or a vote for type B. The voting is done independently. In other words, the two participants have no way to communicate. The following situations can occur:

- If both vote for A, then the third participant will be assigned type A. The group then comprises two type A participants and one type B participant.
- If both vote for B, then the third participant will be assigned type B. The group then comprises two type B participants and one type A participant.
- If one votes for A and one for B, then the type of the third participant will be determined randomly. It is equally likely that type A or type B will be drawn.

Stage 3a: The subsequent decisions on the contributions are not made by all participants at the same time. The third participant is the first-mover. He is the first to make his contribution decision.

Stage 3b: The other two participants in the group then decide on their respective contribution decision.

Stage 4a: After entering the contribution decision, the first-mover is asked of his assessment, what contributions to the project he expects from the participant of type A and from the participant of type B.

Stage 4b: The other two participants are asked about their assessments regarding the contribution of the respective another participant for all possible contribution decisions of the first-mover.

Then the first part of the study ends.

Comprehension questions:

Please answer the comprehension questions on the monitor. These should make it easier for you to familiarize yourself with the decision situation. The following table will help you to answer comprehension questions.

	Income from the project	
Sum of contributions to the project	Type A	Type B
20	10	15
40	20	30
60	30	45

To help you make decisions during the study, you will get a table of the following form:

If you still have questions, please hold a hand out of the cabin - we will come to you.

Contribution of the first-mover (participant 3)	Contribution of type B participant	Contribution of type A participant	Income of the first-mover of		Income of type B participant	Income of type A participant
			type B	type A		
0	0	0	20	20	20	20
...	...	...	...	...	...	...
20	20	20	45	30	45	30

II.A.2.2 Additional Explanations for the First Part of the Study

In each group, the third participant who has been assigned a type by the group by choice, makes his or her contribution decision first. This happens before this participant knows the outcome of the election. The contribution decision is therefore made by the third participant depending on the outcome of the election. For better understanding, look at the following screenshot (Figure II.A.1), which you will see later on the monitor, if you first have to decide on your contribution decision:

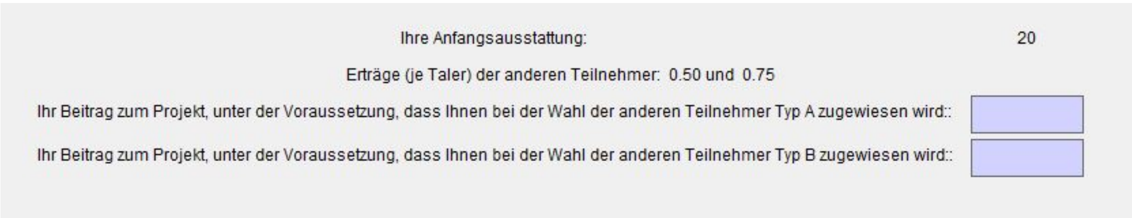


Figure II.A.1: Screenshot with contribution fields for the first-mover in German (endogenous returns)

In the upper field you must enter the number of Talers **you** want to contribute to the project, provided that you are assigned **type A** by the choice of other participants.

In the lower field you must enter the number of Talers **you** want to contribute to the project, provided that you are assigned **type B** by the choice of other participants.

You can enter any integer from 0 to 20 in each field.

When you make your two contribution decisions, you still do not know if you will be type A or type B. You must therefore carefully consider both types of contribution decisions, as both can be relevant for you.

Before the decision of the first participant is communicated to the other two participants, they must complete a contribution table. In the contribution table, you have to decide for each of the 21 possible contribution decisions of the first participant how many Talers you want to contribute to the project yourself. For better understanding, look at the screenshot in Figure II.A.2 below, which you will see later on the monitor in case you are *not* the first to decide on contribution:

The numbers to the left of the blue input fields (on the screenshot) are the possible contributions of the participant who first defined his or her contribution decision. You

Beitragsentscheidungen des ersten Teilnehmers	Ihr Beitrag zum Projekt	Beitragsentscheidungen des ersten Teilnehmers	Ihr Beitrag zum Projekt	Beitragsentscheidungen des ersten Teilnehmers	Ihr Beitrag zum Projekt
0		7		14	
1		8		15	
2		9		16	
3		10		17	
4		11		18	
5		12		19	
6		13		20	

Figure II.A.2: Screenshot with contribution table in German (endogenous returns)

now simply have to enter in each input field how many Talers *you* want to contribute to the project, provided that the first participant provides the specified contribution. **You have to make an entry in each field.**

For example, you must specify how much you contribute to the project when the first participant contributes 0 Talers to the project; how much you contribute when the first participant contributes 1 Taler, or 2 or 3, and so on. You can enter integers from 0 to 20 in each field for **all** fields. Since you do not know at the time of the 21 decisions which of them will be relevant, you should take each one of the decisions carefully.

This decisions and the decisions of the other two participants for this situation will finally determine the payoffs that participants will receive at the end of the first part of the study.

After entering your contribution decisions, you will be asked for your estimates. For better understanding, look at the following screenshot in Figure II.A.3, which you will see later on the monitor when you have to make your assessments first:

In the first two boxes, you will be asked to state your assessments of contribution decisions of **type A participant**, provided that you were assigned type A or type B, correspondingly.

In the next two boxes, you will be asked to state your assessments of contribution decisions of **type B participant**, provided that you were assigned type A or type B, correspondingly.

Similarly, the two participants, whose types were randomly determined, will be asked to state their assessments of the contribution decision of another such participant. For better understanding, look at the following screenshot in Figure II.A.4, which you will see

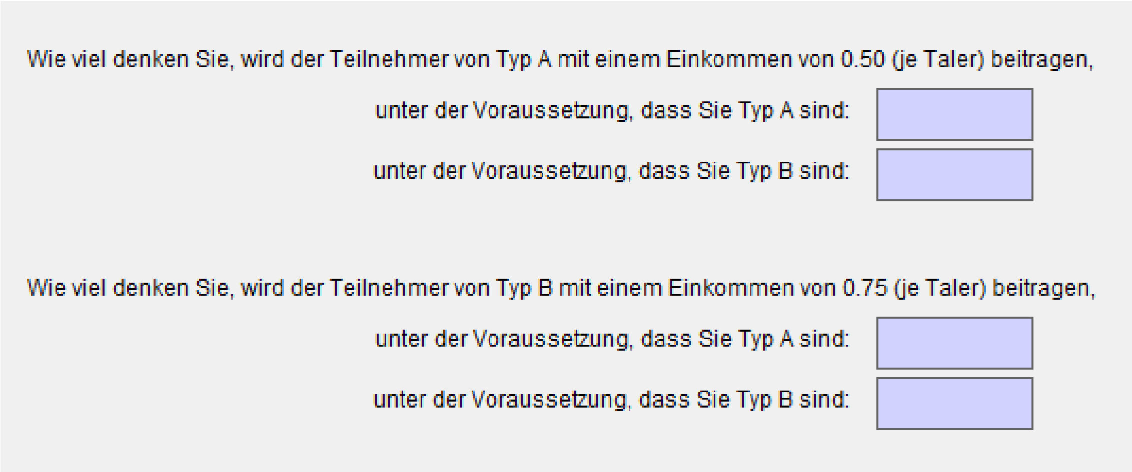


Figure II.A.3: Screenshot with assessments by the first-mover in German (endogenous returns)

later on the monitor if your type was *not determined by the voting*:

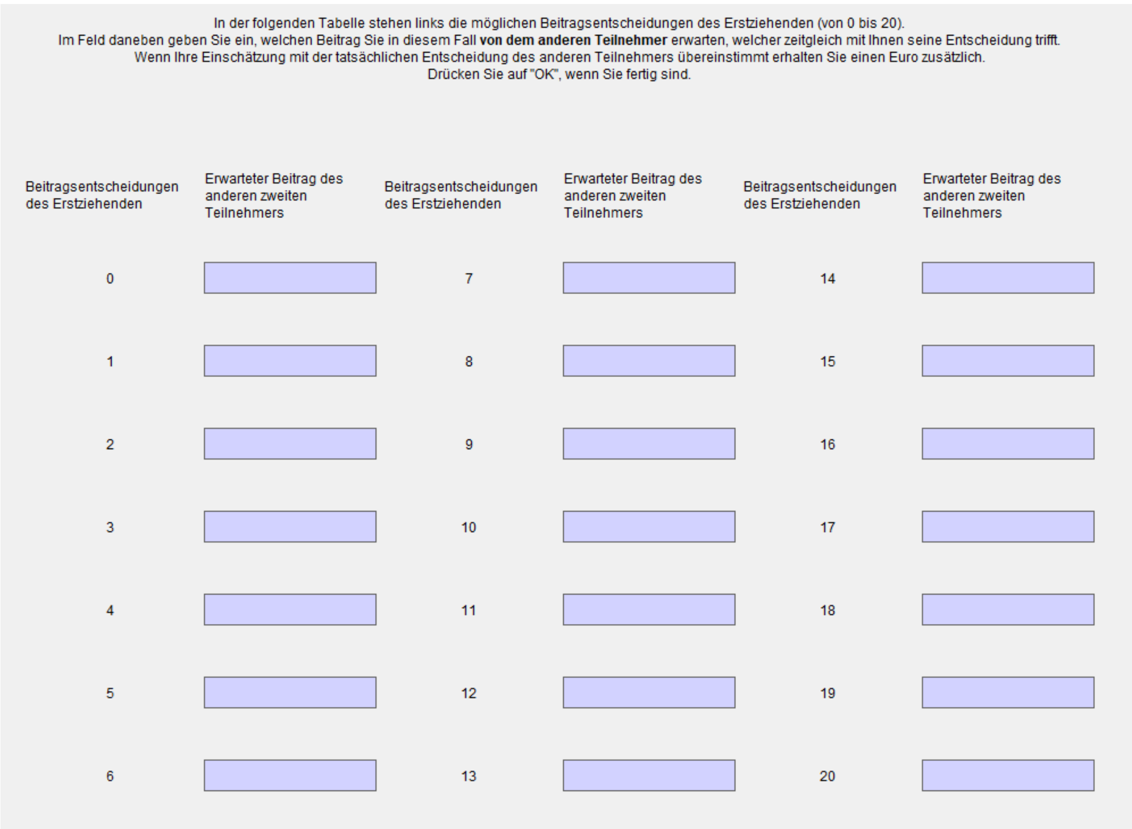


Figure II.A.4: Screenshot with assessments table in German (endogenous returns)

The numbers to the left of the blue input fields (on the screenshot) are the possible contributions of the participant who first defined his or her contribution decision. **You** now need to enter your **assessments** of *the contributions of another participant* for each possible contribution decision made by the first-mover. **You have to make an entry in**

each field.

For example, you need to give your estimates of how much the second participant contributes to the project, if the first-mover contributes 0 Talers to the project, your estimates of how much the second participant contributes to the project, if the first-mover contributes 1 Talers to the project, or 2, 3, etc. You can enter **any** integer from 0 to 20 in each field. You will receive an additional payment of 1 EUR if your assessment matches the actual contribution decision.

Only at the end of the study the first-mover will learn which type was assigned to him. Similarly, the other two participants will learn only at the end of the study, which type was the first-mover assigned and what was his actual contribution decision.

Do you have more questions? If so, please hold a hand out of the cabin.

II.A.3 Instructions for Part 2 (from the Computer Screen)

II.A.3.1 General Information About the Second Part of the Study

The second part of the study comprises of a total of *ten periods*. The decision situation in each period is similar to the previous situation:

You are back in the same trio as before. One participant was randomly determined to be type A, and one other participant was randomly determined to be type B. They will be of the same type as before. At the beginning, if a type has been assigned to you randomly, you will vote for the type of the third player. This player will keep the subsequently assigned type until the end of the study.

Each participant receives 20 Talers at the beginning of *each* period. Your job is again, to decide how many Talers you want to contribute to a project (the rest remains back in your private account).

The contribution decisions are not made by all participants at the same time. Once again, the third participant (first-mover) in your group (the same group member, as in the first part) makes his or her contribution decision first. **Unlike before**, the first-mover **learns which type has been assigned to him prior to his decision**, and thereafter enters only one contribution decision, which immediately becomes payoff-relevant. **Unlike before**, the other two participants do not have to enter their decisions in the form of a contribution table. Instead, **they first learn what type the first-mover has been assigned** and **the first-mover's actual decision**. Thus, they also enter only one contribution decision, which immediately becomes payoff-relevant.

At the end of each period, you will learn about the contributions and pay-outs of all participants in your group. Then the next period begins. Your type and composition of the group remains the same over all ten periods. Likewise, the same participant is the first to enter his or her contribution decision.

The amount of pay-outs per period is determined as before. In the end, you will receive the sum of your pay-outs over all ten periods, with the exchange rate from Euro to Talers now: 1 Taler = 0.03 Euro.

If you have any questions, please hold a hand out of the cabin, and we will come to you.

II.B Additional Regression Results

II.B.1 Additional Results for a Random Effects Model for Contributions in a One-Shot Game

Table II.B.1 shows the comparison of contribution patterns between groups of followers with different returns. Columns 1 to 4 correspond with the estimates on the sub-sample of H-followers, while columns 5 to 8 correspond with the estimates on the sub-sample of L-followers. As in the previous tables, columns 1 and 5 show results for the simplest model (without beliefs and additional controls), models in columns 2 and 6 add beliefs and corresponding cross-terms as explanatory variables, columns 3 and 7 add voting decisions, and columns 4 and 8 include additional individual controls from the questionnaire.

The results presented in Table II.B.1 show substantial differences in contribution patterns of followers with identical returns between groups. First, once beliefs about the contribution of another follower in the group are included (columns 2 and 6), it appears that H-followers follow leaders more closely than L-followers, holding the rest equal. Second, voting for the leader to have higher return from the public good (columns 3 and 7) increases contributions of H-followers, but not L-followers. The inclusion of individual controls (columns 4 and 8) does not change the results. Overall, beliefs about another follower's contribution seem to be more important for followers' contribution decisions than anything else.

Table II.B.1: Estimates for random effects model of conditional contributions by followers (endogenous returns)

	(1) H-followers	(2) H-followers	(3) H-followers	(4) H-followers	(5) L-followers	(6) L-followers	(7) L-followers	(8) L-followers
Leader's contribution	0.486**	0.297*	0.296*	0.297*	0.343***	0.131	0.131	0.134
	(0.157)	(0.120)	(0.120)	(0.120)	(0.102)	(0.093)	(0.093)	(0.094)
H-leader	-1.261	-0.713	-0.715	-0.716	0.377	0.449	0.449	0.448
	(1.107)	(0.714)	(0.715)	(0.719)	(0.297)	(0.331)	(0.331)	(0.333)
H-leader × Leader's contribution	0.060	0.191	0.191	0.191	-0.085	-0.048	-0.048	-0.048
	(0.122)	(0.135)	(0.135)	(0.135)	(0.062)	(0.028)	(0.028)	(0.028)
Belief		0.513***	0.514***	0.513***		0.430**	0.431**	0.425*
		(0.112)	(0.112)	(0.112)		(0.166)	(0.166)	(0.169)
H-leader × Belief		-0.273*	-0.273*	-0.272*		-0.104	-0.104	-0.104
		(0.120)	(0.121)	(0.121)		(0.085)	(0.085)	(0.086)
VotedHigh			2.428	4.748			-1.423	3.076
			(2.526)	(3.511)			(1.706)	(3.790)
Constant	2.029	0.673	-1.150	1.467	1.329	1.138	2.382	-2.675
	(1.196)	(0.693)	(2.313)	(5.316)	(1.307)	(1.144)	(1.303)	(5.250)
Other controls	No	No	No	Yes	No	No	No	Yes
Observations	672	672	672	672	672	672	672	672
Number of individuals	16	16	16	16	16	16	16	16

Notes: *L-follower* is a dummy equal to 1 if the follower has a low return from the public good; *H-leader* is a dummy equal to 1 if the leader in the group was selected to have a high return from the public good; *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader; *VotedHigh* is a dummy equal to 1 if the follower voted for the leader to have a high return. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance.

II.B.2 Additional Results for a Fixed Effects Model for Contributions in a One-Shot Game

Table II.B.2: Estimates for a fixed effects model of conditional contributions (endogenous returns)

	(1) Contribution	(2) Contribution
Leader's contribution	0.532***	0.320**
	(0.138)	(0.116)
L-follower × Leader's contribution	-0.203	-0.201
	(0.167)	(0.148)
H-leader × Leader's contribution	-0.032	0.147
	(0.049)	(0.100)
L-follower × H-leader × Leader's contribution	-0.025	-0.162
	(0.064)	(0.102)
Belief		0.517***
		(0.112)
H-leader × Belief		-0.287*
		(0.128)
L-follower × Belief		-0.096
		(0.201)
L-follower × H-leader × Belief		0.185
		(0.153)
Constant	1.458	0.847
	(0.798)	(0.699)
Observations	1344	1344
Number of individuals	32	32
Fixed Effects	0.34	0.46
model	fe	fe

Notes: *L-follower* is a dummy equal to 1 if the follower has a low return from the public good; *H-leader* is a dummy equal to 1 if the leader in the group was selected to a have high return from the public good; *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader.

Table II.B.3: Estimates for a fixed effects model of conditional contributions by follower (endogenous returns)

	(1) H-followers	(2) H-followers	(3) L-followers	(4) L-followers
Leader's contribution	0.532**	0.320*	0.329**	0.119
	(0.140)	(0.118)	(0.096)	(0.094)
H-leader × Leader's contribution	-0.032	0.147	-0.058	-0.016
	(0.050)	(0.101)	(0.041)	(0.023)
Belief		0.517***		0.422*
		(0.114)		(0.169)
H-leader × Belief		-0.287*		-0.102
		(0.131)		(0.084)
Constant	1.399	0.329	1.517	1.365
	(1.341)	(1.131)	(0.914)	(0.862)
Observations	672	672	672	672
Number of individuals	16	16	16	16
Fixed Effects	0.36	0.48	0.29	0.42
model	fe	fe	fe	fe

Notes: *H-leader* is a dummy equal to 1 if the leader in the group was selected to a have high return from the public good; *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader.

Table II.B.4: Estimates for a fixed effects model of conditional contributions w.r.t. to contributions by the leader with similar/different return (endogenous returns)

	(1) H-followers	(2) L-followers
Contribution by leader with same return	0.466**	0.119
	(0.130)	(0.094)
Contribution by leader with different return	0.320*	0.103
	(0.118)	(0.096)
Belief, when leader has same return	0.230	0.422*
	(0.141)	(0.169)
Belief, when leader has different return	0.517***	0.320*
	(0.114)	(0.145)
Constant	0.329	1.365
	(1.131)	(0.862)
Observations	672	672
Number of individuals	16	16
Fixed Effects	0.48	0.42
model	fe	fe

Notes: *Belief* is a belief from a conditional beliefs table about how much another follower in the group contributes in response to a particular contribution by the leader.

II.C Summary Statistics for Repeated Games

Table II.C.1: Summary statistics for the outcomes in the repeated game (endogenous returns)

Player	**Stat**	**Periods**				**Overall**
		1	**2–5**	**6–9**	**10**	
Contributions						
H-HL, leader	mean	15.8	13	11.1	12.7	12.5
	std	(5.88)	(7.15)	(8.14)	(8.31)	(7.62)
	median	[20]	[15]	[10]	[17]	[15]
H-HL, L-follower	mean	12.4	9.46	6.13	7.46	8.22
	std	(7.84)	(8.83)	(8.13)	(9.06)	(8.63)
	median	[15]	[7.5]	[2]	[3]	[5]
H-HL, H-follower	mean	11.2	11.7	10.8	3.92	10.5
	std	(8.43)	(7.66)	(8.35)	(6.78)	(8.16)
	median	[10]	[11]	[10]	[0]	[10]
L-HL, leader	mean	20	7.67	6.5	0	7.67
	std	(0)	(7.24)	(6.49)	(0)	(7.63)
	median	[20]	[10]	[7.5]	[0]	[10]
L-HL, L-follower	mean	10	8.75	3.83	0	6.03
	std	(10)	(6.59)	(6.25)	(0)	(6.98)
	median	[10]	[10.5]	[0]	[0]	[2.5]
L-HL, H-follower	mean	12.7	9.25	7.58	0.333	8.03
	std	(2.52)	(6.92)	(8.55)	(0.577)	(7.44)
	median	[13]	[11]	[4.5]	[0]	[9.5]
Payoffs						
H-HL, leader	mean	33.7	32.6	29.9	25.4	30.9
	std	(6.82)	(9.68)	(9.5)	(6.34)	(9.29)
	median	[31]	[30]	[27.8]	[25.8]	[30]
H-HL, L-follower	mean	27.3	27.6	27.9	24.6	27.4
	std	(7.49)	(5.39)	(6.11)	(7.73)	(6.16)
	median	[24]	[29]	[27.5]	[22.5]	[27.5]
H-HL, H-follower	mean	38.3	34	30.2	34.1	32.9
	std	(8.2)	(9.47)	(9.43)	(10.4)	(9.66)
	median	[36.2]	[31.5]	[28.5]	[33.8]	[31.2]
L-HL, leader	mean	21.3	25.2	22.5	20.2	23.2
	std	(4.65)	(3.15)	(4.79)	(0.289)	(4.14)
	median	[20]	[25.5]	[20]	[20]	[21.5]
L-HL, L-follower	mean	31.3	24.1	25.1	20.2	24.8
	std	(5.62)	(3.98)	(5.95)	(0.289)	(5.34)
	median	[30]	[25.5]	[25]	[20]	[25.5]
L-HL, H-follower	mean	39.3	30	25.9	19.9	28.3
	std	(7.77)	(9.15)	(6.17)	(0.144)	(8.59)
	median	[40]	[32.6]	[24.4]	[20]	[25.1]

Chapter III

On the Comparison of the Effect of Leading-by-Example in Asymmetric Public Goods Games with and without Endogeneity in the Composition of Returns*

III.1 Introduction

Over the past decades, economists have suggested numerous mechanisms to overcome the free-rider problem in social dilemmas. One of the most prominent topics in this area is the role of institutions in creating incentives for collaboration, particularly their successful implementation when the establishment of these institutions is left upon the involved participants. The role of such endogenous institutions in public goods games has been widely studied in symmetric groups, where group members are identical and equal contributions, thus, lead to equal payoffs. However, in natural environments, it is often difficult to assume that group members are identical for the groups to be symmetric. With the presence of asymmetries between group members, the possibility for a conflict of interests to emerge increases. Consequently, policy-makers have to take into account yet another dimension in the parameters of the institution design so that the institution works as intended, i.e. creates incentives for cooperation that are robust to existing asymmetries. Nevertheless, currently there is only limited evidence of the effect of heterogeneity among group members on the success of certain institutions, although the functioning of these institutions has been widely studied.

I study leading-by-example in the public goods game with asymmetric individual returns. Leading-by-example is an institution that transforms a standard public goods game into a sequential one: one player moves first (the leader), and other players (the followers) move after observing the leader's contribution. Below, I explore (i) how the functioning of this institution varies when the composition of returns changes, and (ii) how it changes

*I would like to thank Sebastian Kube and Sebastian Schaube for their insightful comments, and the participants of the Bonn Applied Micro Workshop, Copenhagen Workshop on Endogenous Institutions in Social Dilemmas, and ESA World Meeting 2018 in Berlin for fruitful discussions.

when the composition of returns is determined endogenously by a sub-group of players rather than being exogenously assigned by the experimenter.

I use two laboratory experiments (presented in chapters I and II) to study cooperation in the provision of public goods among groups of three players who differ in their marginal per capita returns from the public good (MPCR), i.e. have high or low returns. The experiments differ in the way the composition of returns in the group is established: in one experiment the composition is predetermined and in another experiment it is established by followers through a voting procedure. Followers always have different returns (high or low) and vote for the return of the leader in their group. The experimental approach allows me to mitigate the selection bias in followers' reactions and measure their expectations about unobserved actions of the other follower, namely to account for all payoff-relevant factors that may influence contribution decisions.

I find that, when leaders' returns are set exogenously, followers cooperate more willingly with leaders who have lower returns from the public good. This result goes in line with previous findings in the literature, namely that players are more likely to cooperate with leaders who are perceived as "disadvantaged" players (see, for example, Glöckner et al., 2011). When leaders' returns are left upon the decision by the second-movers, followers' responses to leaders' contributions are no longer statistically different. The majority of followers vote for the leader to have large benefits from the public good, and leaders do reciprocate being elected to have high returns with increased contributions. Such contrasting difference in followers reactions between groups with leaders having exogenous or endogenous returns suggests that endogenizing the compositions of returns may have spillover effects on the provision of public goods. The effectiveness of leading-by-example, as an institution meant to function as a coordination device, changes, depending on the way leaders' returns are determined. My findings suggest that players in groups with predetermined composition of returns who have low-benefiting leaders may better overcome problems associated with inequality considerations and achieve efficiency more easily.

Going deeper into understanding why there occurs such a difference in followers' responses between groups with leaders having exogenous and endogenous returns, I find that responses to leaders' contributions fall among followers with low returns who vote for leaders to have high returns: they seem to ignore the leader's contributions and contribute much smaller amounts. Only those low-benefiting followers who vote for their leaders to have low returns also demonstrate high contribution responses. This result suggests that voting and contribution decisions are not completely independent and may jointly reflect players' preferences regarding the provision of public goods.

I also investigate how the patterns in conditional contributions translate into dynamics by letting subjects play a repeated version of the game. The results suggest that sustaining cooperation over time is very sensitive to the endogeneity of the composition of returns: contributions in groups with low-benefiting leaders fall more rapidly when leaders' returns are decided by followers but remain high over the whole course of the repeated interaction when returns are set exogenously. Surprisingly, leaders elected to have low returns cannot be blamed for poor cooperation dynamics: they contribute the full endowment in the first round of the repeated game, most likely due to strategic considerations. However, followers in such groups did not reciprocate leaders' high initial contributions, so overall contributions decrease in subsequent periods.

The rest of the chapter is organized as follows. I explain the design of the two laboratory experiments in Section III.2. Section III.3 elaborates on the key predictions from theoretical models of social preferences to explain what differences one can expect. I present the patterns in followers' contribution decisions obtained from the two economic experiments and describe main differences among them in Section III.4. Section III.5 provides the results from the regression analysis for conditional contributions. Section III.6 describes the differences in the way contribution preferences translate into dynamics in the two experiments. Section III.7 concludes.

III.2 The Two Experiments

In my experiments, subjects play a public goods game: they decide how much of their endowment (of 20 points) to keep and how much to contribute to the public good. The game is played in groups of three players and contributions are made sequentially: a randomly-selected player makes the decision first (the leader), the other two (the followers) decide upon observing the leader's contribution. There are two types of MPCR, high and low. One follower is randomly allocated to always have a high return (H-follower) and the other follower has a low return (L-follower).

Leaders' returns can either be high or low: in the first experiment, called Exo treatments, leaders' returns are predetermined. Half of the groups have leaders with high returns and the other half have leaders with low ones. Therefore, I have two types of groups: there are either two players with high returns and one player with a low one, or two players with low returns and one player with a high one. As followers in the same group always have different returns, groups then vary according to the leaders' returns, high and low, and are called then H-HL groups (with H-leaders) and L-HL groups (with L-leaders), correspondingly. In the second experiment, called the Endo treatment, leaders' returns are chosen by the followers through a voting procedure before making contribution decisions.

If leaders with high returns make different contribution decisions than leaders with low returns, it would be difficult to directly compare followers' responses to leaders' contributions.[1] In order to solve this problem and isolate followers' contributions from their prior beliefs about leaders' contributions, I use a strategy method (Selten, 1967). I measure contribution preferences, namely contribution decisions in response to all possible leaders' contributions (21 decisions per follower, as contributions are only allowed to be natural numbers or zero).

[1] Because the supports of the distributions of followers' contributions differ in this case.

As followers condition their own contribution decision on the leader's contribution, the only remaining factor that can potentially affect the payoff of a follower is the contribution by the other follower. For this reason, I elicit followers' conditional beliefs:[2]how much followers expect the other follower in their group to contribute, conditional on the leaders' decision.

In the second experiment, the Endo treatment, I elicit two sets of followers' contribution preferences and conditional beliefs: in case the leader has a high return and in case the leader has a low one. Leaders in the Endo treatment also make two contribution decisions, conditional on having potentially high or low returns. The payoffs are determined according to voting and contribution decisions of players at the end of experiments.

With the measures for all payoff-relevant factors described above, the observed outcomes may vary in case players have in mind different models of social comparison that might affect their contribution decisions. In the next section, I discuss what patterns in contribution behavior one may expect.

III.3 Predicted Differences

Different models of social preferences predict different equilibria in a public goods game. Rather than using a particular model of social preferences, I study the behavior in a public goods game with respect to contribution norms that correspond with one of the equilibrium allocations predicted by these models. According to Reuben and Riedl (2013), two normatively appealing contribution rules can potentially arise in groups with heterogeneous returns: the equality of contributions and the equality of payoffs.[3] Cappelen et al. (2007) proposed a framework how people select between different "fairness ideals". Using their approach with the two selected contribution rules, it is easy to show that players with high returns prefer the equality-of-contributions rule and players with low returns prefer the equality-of-payoffs rule (see also Neitzel and Sääksvuori, 2013).[4] Given the sequential moves in my game, it follows that players in H-HL groups should contribute according to equal payoffs, and players in L-HL groups should contribute equally since the leader can choose her preferred contribution rule and signal it to the other players via her contribution. This prediction immediately implies that in Exo treatments followers in L-HL groups should contribute more generously than followers in H-HL groups, given the same level of the leader's contribution.

Under the assumption that H-players prefer the equality-of-contributions rule and L-players prefer the equality-of-payoffs rule, one can show that L-followers should always vote for the leader to have a high return since their expected payoffs are then higher.[5] H-

[2]Following Fischbacher and Gächter (2010), I call players' expectations *players' beliefs*.

[3]These rules apply for groups that do not achieve efficiency in the provision of the public good, i.e. do not contribute full endowments.

[4]Reuben and Riedl (2013) find that these are the two rules players with heterogeneous returns tend to enforce in a public goods game with a peer-to-peer sanctioning.

[5]Given the leader's contribution of x, the L-follower's payoff (under the patterns described above) is $20+0.5x$ in L-HL, and is $20+0.8x$ in H-HL (here, I assume that the followers take the leader's contribution

followers should vote for H-leaders only if they expect L-leaders to contribute substantially less than H-leaders (otherwise, their expected payoffs are higher with L-leaders).[6] These expectations are not unfounded. H-leaders might contribute large amounts because they might want to cooperate with H-followers, since it is profitable for these two players to cooperate even if the L-follower free-rides. However, in the Endo treatment, H-leaders may also contribute more because they reciprocate being elected to receive high returns. Therefore, we might expect to see H-HL groups more often in the Endo treatment, with L-followers voting for H-leaders more often than H-followers.

There might be a strong dependence between voting and contribution decisions in the Endo treatment (see Figure III.1). The conditional contribution tables, in fact, imply that followers know what return the leader has for a particular contribution table they fill out. Therefore, they know the "outcome" of the voting procedure for this case and can compare it with their own voting decision.[7] Follower 1 might want to follow the leader only if the leader has the return for which follower 1 voted: if follower 1 voted for the leader to have a high return and the leader received a low one, it clearly indicates that the other follower voted for the low return. This might change follower 1's responses to L-leaders' contributions depending on his own voting decision and the observed leader's return.[8] This problem does not allow drawing sharp predictions about the differences in contribution patterns between H-HL and L-HL groups in the Endo treatment.[9]

To sum up, I expect the following differences to be observed in the two experiments.

1. H-leaders contribute in the Endo treatment more than in the Exo treatment;
2. H-HL groups are more likely to emerge in the Endo treatment, as followers should expect higher contributions by H-leaders;
3. Free-riding is more pronounced among followers in H-HL groups in the Endo treatment than in Exo treatments;
4. In Exo treatments, followers in L-HL groups should contribute more generously than followers in H-HL groups, given the same level of the leader's contribution; it is difficult to predict differences in contributions between groups with H- and L-leaders in the Endo treatment.

as a reference). If H-leaders contribute more than L-leaders, L-followers are even more better off.

[6] Given the leader's contribution of x, the H-follower's payoff (under the patterns described above) is $20+1.25x$ in L-HL, and is $20+0.8x$ in H-HL (here, I assume that the followers take the leader's contribution as a reference). Therefore, for H-followers to prefer H-leaders, H-leaders should contribute 56.25% more than L-leaders.

[7] The voting does not necessarily need to happen already: one can consider voting as a signal for the preferred leaders' return, so followers should already know what signal they want to send.

[8] If follower 1 voted H but got an L-leader, he might be disappointed and contribute less than he would contribute in case he voted L and got an L-leader. It is unclear whether one should expect a similar effect for those who vote L but get H-leaders.

[9] In Exo treatments, one might expect efficient allocations to occur in L-HL groups since for L-players it does not bring any monetary costs if they contribute the full endowment (compared with contributions that lead to equal payoffs), only differences in payoffs with the H-follower. In the Endo treatment, it is unlikely to observe such efficient allocations, first, because L-HL groups are less likely to emerge and, second, because voting decisions might interfere with contribution decisions.

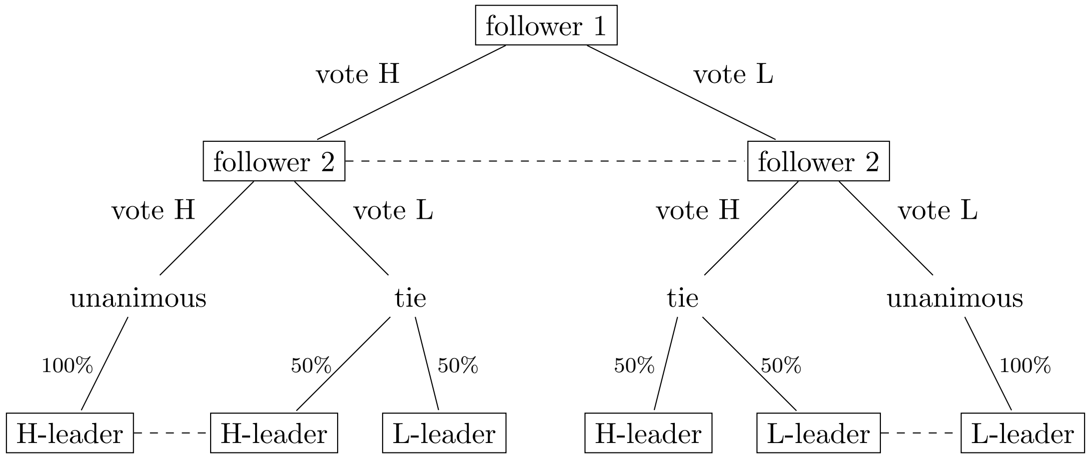


Figure III.1: Voting decisions as an extensive-form game

III.3.1 Procedures

The instructions (see Appendices in chapters I and II) explained to participants the payment system in a public good problem and the differences between players. After participants had read the instructions, they had to answer control questions to ensure their understanding of payoff structure and the dilemma situation. The experiment started only after all participants answered all questions correctly. Each player was endowed with $y = 20$ points and allocated these points between investments in the public good and private account. I use neutral wording in the instructions and during the experiment (e.g. players with low and with high returns are referred to as players of type A and of type B, respectively, with known MPCR).

After playing the one-shot game, participants were asked to play the repeated version of a public goods game. These games were designed to study how contribution preferences transform into dynamics. I report the results from this part of the experiment after the results from the main treatment. The repeated games are played for ten rounds in partner matching mode.[10] In order to isolate the decisions in a repeated game from the decision in the main treatments, I report the outcomes of the one-shot game only at the end of the experiment. At the end of each period of repeated games, players observe the actual contributions and payoffs of *all* players in their group.

Groups in the repeated games comprised of the same participants as in the one-shot game. Followers kept their returns same as in a one-shot game, as well as leaders in Exo treatments, but in the Endo treatment followers had to vote on the leader's return again before playing the repeated game, albeit only once. Unlike previously, all players learn the outcome of voting before they make contribution decisions in the Endo treatment. Additionally, in Exo treatments, groups were randomly allocated to two additional groups of treatments: half of L-HL groups and half of H-HL groups played a repeated sequential

[10] Despite partner matching mode in a repeated game allows for the strategic behavior on the one hand, on the other hand it allows excluding unobserved heterogeneity in beliefs and cooperative preferences that occurs if one uses stranger matching mode, and it allows for a more precise updating of beliefs.

public goods game, and the other half played a repeated simultaneous game. These additional treatments are necessary to make the comparison of contributions and outcomes between simultaneous and sequential game, because it is not possible to properly elicit contribution preferences in a simultaneous game which can be compared with those in a repeated game. The repeated game played after the Endo treatment was a repeated sequential one. In a repeated sequential game, in each round the leader decides on his contribution first, then the followers actually observe his decision and make contributions simultaneously. In a simultaneous game all players make simultaneous contributions.

At the end of the experiment participants were asked to complete a questionnaire which included questions on gender, age, field of study, self-assessment of risk preferences, trust and power distance (see House et al., 2004).

The experiments took place in the Bonn Econ Lab in July 2017 and February 2018. In total, I have 93 participants for Exo treatments (45 for H-HL and 48 for L-HL) and 48 participants for the Endo treatment. All participants were recruited using HROOT recruitment software (Bock et al., 2014). The participants were mostly students from various disciplines from the University of Bonn. All treatments were programmed using the z-Tree software (Fischbacher, 2007).

In Exo treatments, participants earned on average 14 EUR (slightly above 16 USD), including a show-up fee of 4 EUR. Each session in Exo treatments lasted roughly 80 minutes (including random allocation of the participants into cubicles, quiz, questionnaire and payments). In the Endo treatment participants earned on average 30.3 EUR (around 37.7 USD) including a show-up fee of 4 EUR.[11] Each session in the Endo treatment lasted around 120 minutes (including random allocation of the participants into cubicles, quiz, questionnaire and payments).

III.4 Results

Voting decisions in the Endo treatment are in line with the predicted behavior: more than 80% of all followers voted for their leaders to have high returns; furthermore, L-followers voted for H-leaders more often than H-followers. Since ties in voting decisions break up randomly, ultimately I have eleven H-HL groups out of sixteen total. Leaders contribute more in the Endo treatment when assigned high returns (on average, they contribute 14.7 points versus 9 in case they have low returns), but not in Exo treatments (both H- and L-leaders contribute 10.9 points on average).

First, I focus on the key differences in followers' contribution preferences (conditional contributions) between groups with exogenous and endogenous composition of returns,

[11] During the experiment subjects earned their payoffs in points. Initially, I set the exchange rate of 1 point = 0.1 EUR for the points earned in a one-shot game, and at a rate of 1 point = 0.02 EUR for repeated games. In order to meet the minimum pay-out requirement of the Bonn Econ Lab, for Endo treatment I recalculated the exchange rate based on cooperation rates from Exo treatments, resulting in an exchange rate of 1 point = 0.6 EUR for the points earned in a one-shot game, and at a rate of 1 point = 0.03 EUR for a repeated game.

and how they relate to followers' expectations about the contribution of the other follower (conditional beliefs).

III.4.1 Patterns of Conditional Contributions

Figure III.2 below shows the patterns of followers' conditional contributions. Contribution patterns in groups with exogenous composition of returns are presented on panels to the left, while patterns in groups with endogenous composition of returns are presented on panels to the right. There are two differences between these two treatments:

1. Conditional contributions of followers differ with respect to the leader's return in Exo treatments (panels III.2a and III.2c), but appear similar in the Endo treatment (panels III.2b and III.2d)
   - in Exo treatments, followers respond more generously to contributions by L-leaders;
2. While followers show nearly identical contribution preferences with respect to H-leaders (solid circles lines), these patterns differ with respect to L-leaders (empty triangles lines) in both Exo and Endo treatments
   - for both L- and H-followers, contributions with respect to L-leaders seem to be less generous in the Endo treatment than in Exo treatments.

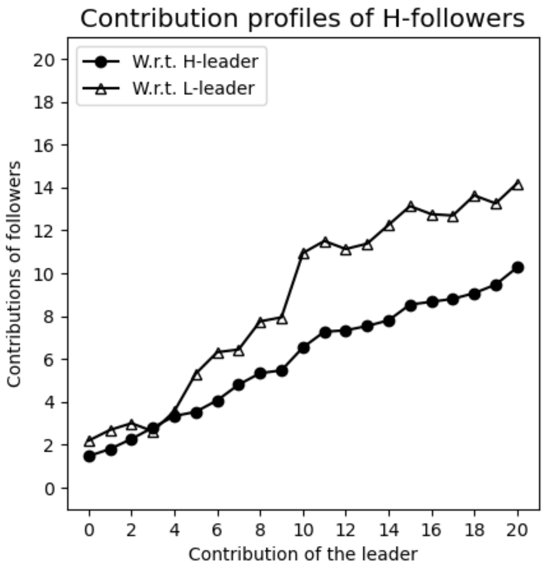
Such asymmetries in contribution patterns may occur for a number of reasons. First, the differences in contribution preferences in Exo treatments may be the result of holding different beliefs about the contribution of the other follower. Second, in addition to heterogeneous beliefs and cooperation preferences, the differences in reactions between Exo and Endo treatments might also stem from the simultaneity of decisions on contribution and voting: followers might condition their contribution preferences on voting outcomes, implicitly threatening the other follower with low contributions if the realized voting outcome is less favorable.[12]

Below I present evidence that followers' beliefs about the contribution of the other follower do not change much when leaders have different returns, but they do vary between treatments, especially for L-followers.

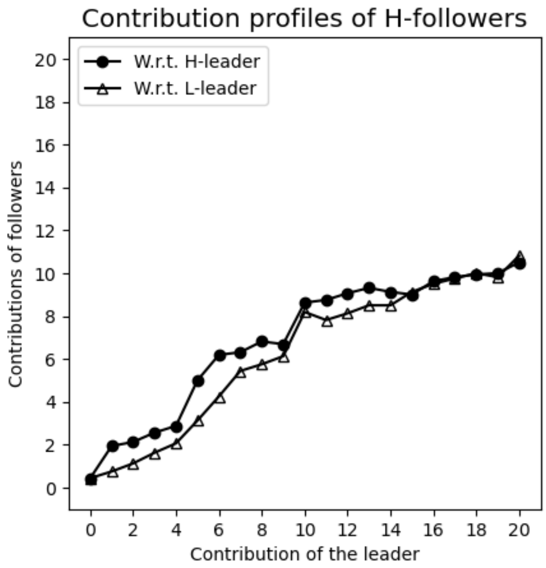
III.4.2 Conditional Contributions and Conditional Beliefs

I contrast conditional contributions and conditional beliefs about the contribution of the other follower in the group. I find that these beliefs are crucial in determining followers' contributions. Analyzing the relation of these conditional beliefs to contribution preferences may help to clarify how followers compare themselves to other players in the group.

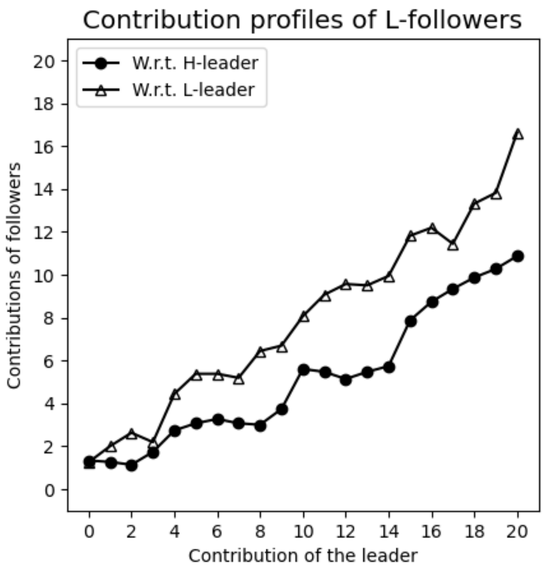
[12] It is also ambiguous whether these between-treatment differences come from the possibility to choose the leader's return or the fact that in Exo treatments I have the *between*-subjects design, while in Endo treatment I have the *within*-subjects design. Unfortunately, I am not aware of the way in which to distinguish these two possibilities.



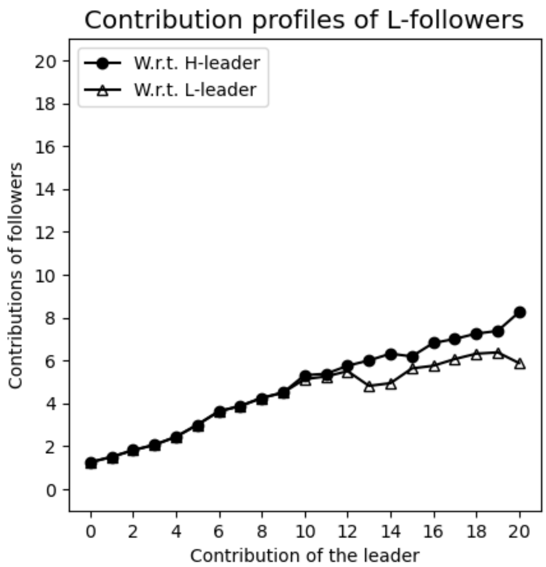
(a) H-followers, Exo treatments



(b) H-followers, Endo treatment



(c) L-followers, Exo treatments

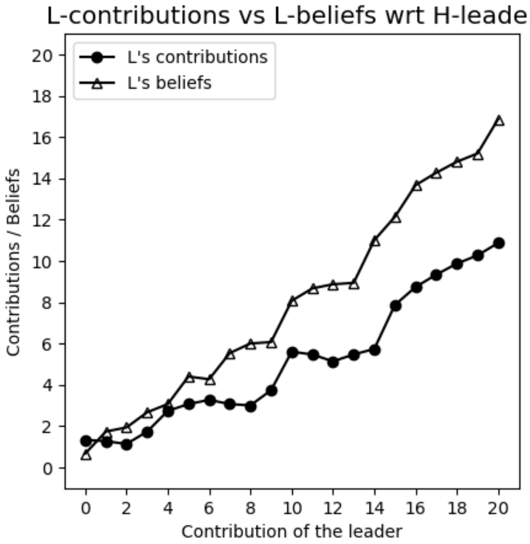


(d) L-followers, Endo treatment

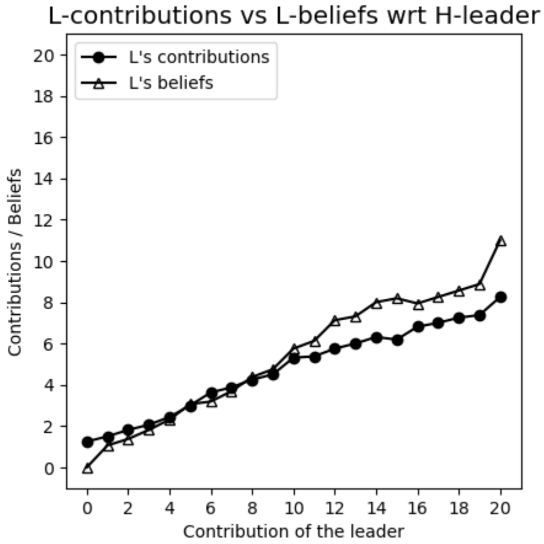
Figure III.2: Averaged conditional contributions with respect to *H-* and *L-leaders*: between-subject (Exo) to the left, and within-subject (Endo) to the right

Figures III.3 and III.4 below plot conditional contributions and conditional beliefs versus corresponding contributions of H- and L-leaders, respectively.

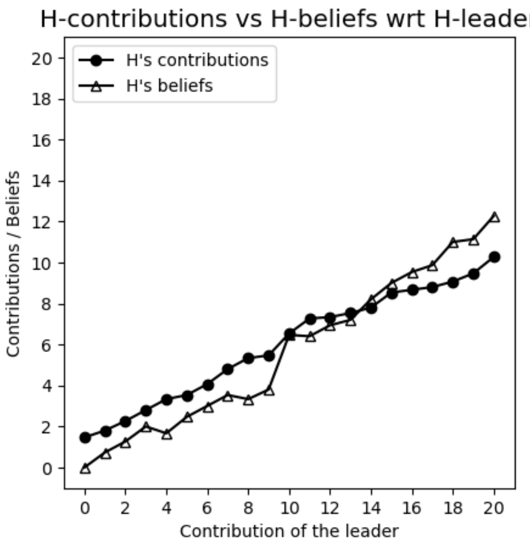
According to the patterns depicted in Figure III.3, followers contribute nearly as much as they expect the other follower in the group to contribute. The only exception is the relation between contributions and expectations of L-followers in H-HL groups (see panels III.3a and III.3b): in Exo treatments, L-followers contribute much less than they expect H-followers to contribute in response to H-leaders' contributions. A similar pattern can also be noticed for the Endo treatment, although the divergence is less pronounced



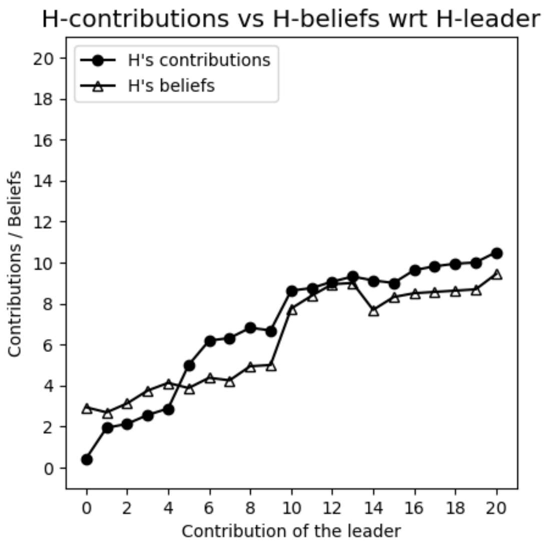
(a) L-followers, Exo treatments



(b) L-followers, Endo treatment



(c) H-followers, Exo treatments



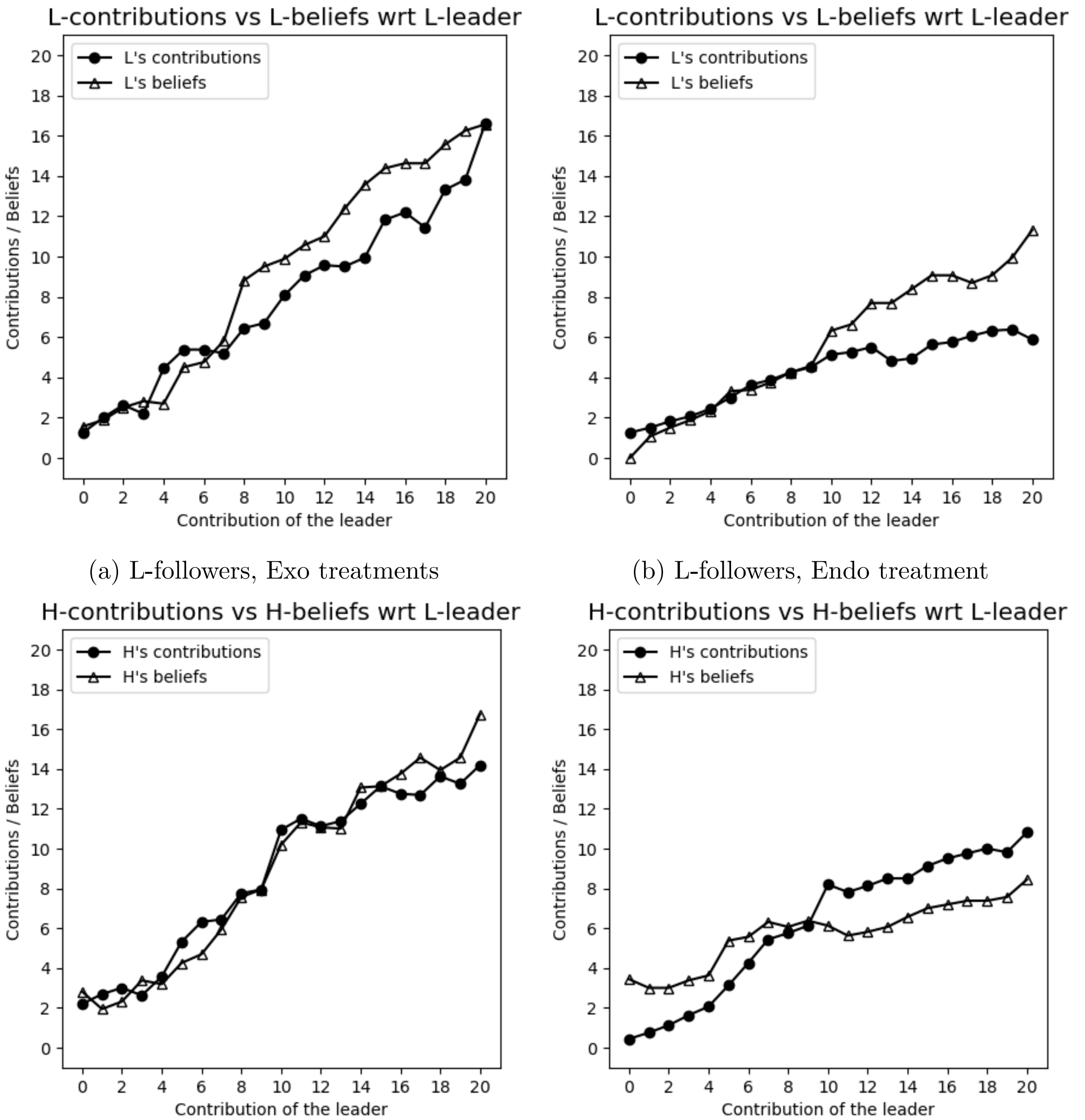
(d) H-followers, Endo treatment

Figure III.3: Averaged followers' conditional contributions and conditional beliefs, with respect to the *H-leader*'s contribution decision

and can only be seen for relatively large leaders' contributions.

By contrast, the patterns depicted in Figure III.4 suggest that there are substantial differences between followers' beliefs and their contributions in the Endo treatment when the leader has a low return (see panels III.4b and III.4b). L-followers contribute much less than they expect H-followers to do when leaders' contributions are sufficiently large. However, H-followers hold unexpectedly high beliefs about the contributions of L-followers: for relatively small leaders' contributions, they expect L-followers to contribute quite large amounts.

Apart from large beliefs of H-followers in L-HL groups in the Endo treatment, all other observations align well with the predicted differences. First, the patterns of con-



(a) L-followers, Exo treatments (b) L-followers, Endo treatment

(c) H-followers, Exo treatments (d) H-followers, Endo treatment

Figure III.4: Averaged followers' conditional contributions and conditional beliefs, with respect to the *L-leader*'s contribution decision

ditional contributions in Exo treatments confirm that followers respond more generously to L-leaders, but not in the Endo treatment. This effect can only partially be attributed to the differences in conditional beliefs about the contribution of the other follower, as the relation between contributions and beliefs changes across groups and between the experiments. Second, in the Endo treatment, followers vote for H-leaders more often, and the leaders that are elected to have high returns contribute more. Third, L-followers do free-ride more in the Endo treatment after they vote for the leader to receive a high return.

In the regression analysis presented in the next section, I analyze the extent to which the differences in conditional contributions are explained by the composition of returns and the ability to vote for the preferred composition, taking into account asymmetries in conditional beliefs.

III.5 Regression Analyses for Conditional Contributions

I perform regression analyses to control for the set of individual characteristics (gender, dummies for high/low risk aversion, high/low trust, and strong/weak power distance) and conditional beliefs. Furthermore, they allow me to ascertain whether differences in contribution patterns result from the asymmetries in returns or stem from the fact that followers are allowed to choose the composition of returns in their group.

I estimate random effects models for conditional contributions on the corresponding leader's contributions,[13] own conditional beliefs about the contribution of the other follower in the group, and individual characteristics. Leaders' contributions and own conditional beliefs are split into two variables, conditional on whether the leader and the follower have the same returns or different ones.[14]

The model was estimated separately for different groups of followers (with high and low returns) and for Exo and Endo treatments, respectively. The estimates are presented in Table III.1 below (coefficients for individual characteristics are not reported).[15] Columns (1) and (4) contain the estimates for H- and L-followers in Exo treatments, while columns (2) and (5) report the estimates for H- and L-followers in the Endo treatment, respectively. The estimates presented in columns (1)–(2) and (4)–(5) were obtained using the same set of explanatory variables for both experiments (i.e. without voting decision of followers in the Endo treatment, they were added in columns (3) and (6)).

There are minor differences in the estimated coefficients for H-followers between the treatments (compare columns (1) and (2)): in the Endo treatment, when the leader also has high returns, followers with high returns seem to rely more on leaders' contributions and less on their beliefs about the contribution of the other follower. L-followers seem to rely mostly on the contribution by the leader with similar returns and not on their expectations in Exo treatments (see column (4)). Contrarily, in the Endo treatment (column (5)), their contributions seem to depend solely on their beliefs.

Voting decisions (a dummy that equals one if the follower voted for the leader to have high returns) seem to increase average contributions of H-followers and decrease average contributions of L-followers (see columns (3) and (6)). However, it is unclear whether voting has an impact on reactions to leaders' contributions. In order to test this, I estimate a similar model with additional interactions of leaders' contributions and followers' voting decisions. The results are presented in Table III.2 below.

[13] For each follower I have 21 contribution entries for 21 possible leader's contributions.

[14] These variables were set to zero if the condition is not satisfied, i.e. for L-(H-)followers, contribution by the leader with the same return equals the corresponding contribution by the L-(H-)leader, and zero otherwise.

[15] The most important factor for contribution decisions is (general) trust: players who strongly agree that one can trust people in general contribute much more than others.

Table III.1: Estimates for random effects model of conditional contributions with respect to contributions by the leader with similar/different returns

	Exo treatment	Endo treatment		Exo treatment	Endo treatment	
	(1) H-followers	(2) H-followers	(3) H-followers	(4) L-followers	(5) L-followers	(6) L-followers
Leader with same return	-0.778	-0.716	-0.711	1.766	-0.448	-0.452
	(2.132)	(0.719)	(0.680)	(1.257)	(0.333)	(0.332)
Contribution by leader with same return	0.085	0.488***	0.210	0.457***	0.134	0.493***
	(0.047)	(0.143)	(0.231)	(0.123)	(0.094)	(0.064)
Contribution by leader with different return	0.320*	0.297*	0.105	-0.052	0.086	0.452**
	(0.125)	(0.120)	(0.196)	(0.079)	(0.097)	(0.142)
Belief, when leader has same return	0.582***	0.241	0.176	0.273	0.425*	0.441**
	(0.156)	(0.139)	(0.152)	(0.171)	(0.169)	(0.164)
Belief, when leader has different return	0.452***	0.513***	0.432***	0.687***	0.322*	0.329*
	(0.125)	(0.112)	(0.118)	(0.107)	(0.145)	(0.140)
VotedHigh		4.748	1.091		3.076	7.234*
		(3.511)	(3.311)		(3.790)	(3.548)
VotedHigh × Leader's Contribution (same return)			0.391			-0.420***
			(0.244)			(0.084)
VotedHigh × Leader's Contribution (different return)			0.296			-0.423**
			(0.251)			(0.146)
Constant	-0.419	1.467	4.912	0.302	-2.227	-5.885
	(2.549)	(5.316)	(5.540)	(2.192)	(5.293)	(4.851)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	651	672	672	651	672	672
Number of individuals	31	16	16	31	16	16

Notes: *Belief* is the belief from a conditional beliefs table about how much the other follower in the group contributes in response to a particular contribution by the leader. *VotedHigh* is a dummy equal to 1 if the follower voted for the leader to have a high return. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance. In the Endo treatment, contribution profiles are analyzed conditional on the leader's return. Standard errors are clustered at the individual level.

Table III.2: Decomposing the effect of voting behavior onto conditional contributions

	(1) H-followers	(2) H-followers	(3) L-followers	(4) L-followers
Leader with same return	-0.716	-0.711	-0.448	-0.452
	(0.719)	(0.680)	(0.333)	(0.332)
Contribution by leader with same return	0.488***	0.210	0.134	0.493***
	(0.143)	(0.231)	(0.094)	(0.064)
Contribution by leader with different return	0.297*	0.105	0.086	0.452**
	(0.120)	(0.196)	(0.097)	(0.142)
Belief, when leader has same return	0.241	0.176	0.425*	0.441**
	(0.139)	(0.152)	(0.169)	(0.164)
Belief, when leader has different return	0.513***	0.432***	0.322*	0.329*
	(0.112)	(0.118)	(0.145)	(0.140)
VotedHigh	4.748	1.091	3.076	7.234*
	(3.511)	(3.311)	(3.790)	(3.548)
VotedHigh × Leader's Contribution (same return)		0.391		-0.420***
		(0.244)		(0.084)
VotedHigh × Leader's Contribution (different return)		0.296		-0.423**
		(0.251)		(0.146)
Constant	1.467	4.912	-2.227	-5.885
	(5.316)	(5.540)	(5.293)	(4.851)
Other controls	Yes	Yes	Yes	Yes
Observations	672	672	672	672
Number of individuals	16	16	16	16

Notes: *Belief* is the belief from a conditional beliefs table about how much the other follower in the group contributes in response to a particular contribution by the leader. *VotedHigh* is a dummy equal to 1 if the follower voted for the leader to have a high return. Additional controls include dummies for gender (female), high/low risk aversion, high/low trust, and high/low power distance. In the Endo treatment contribution profiles are analyzed conditional on the leader's return. Standard errors are clustered at the individual level.

The estimates reported in Table III.2 provide an important insight: while contribution decisions of H-followers do not seem to reflect their voting decisions, the voting decisions of L-followers clearly shape their contributions.[16] Followers with low returns who *did not* vote for the leader to have high returns follow their leaders closely, independently on the leader's return. At the same time, L-followers who voted for H-leaders seem to ignore leaders' contributions and condition their contribution decision on their own expectations about the contribution of the other follower.

III.6 Contribution Dynamics

I let subjects in my experiments play a repeated sequential version of the game to study how contribution preferences translate into dynamics. The repeated games were played for ten consecutive rounds after the elicitation of conditional contributions and beliefs but before reporting the outcomes of the one-shot game. In the Endo treatment, followers were asked to vote for the return of the leader in their group once at the start of the repeated game. Voting outcomes for the repeated game now become available to all players before they have to make contribution decisions.

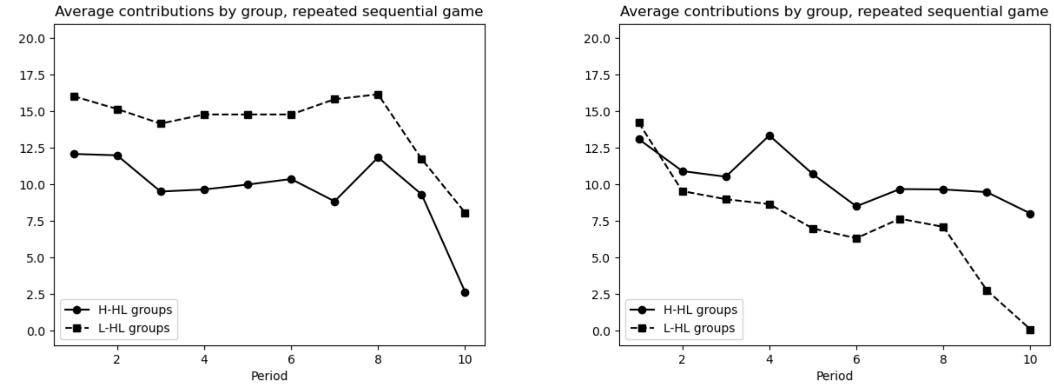
Almost all of the differences in conditional contributions that we observed in the one-shot game preserve in a dynamic setup: H-leaders are elected more often in the Endo treatment, and groups that chose leaders to have low returns demonstrate weaker cooperation.

In the repeated sequential games in the Endo treatment, voting decisions were identical to those in the one-shot game, i.e. more than 80% of followers voted to have H-leaders; only three subjects voted for the leader to have low returns in both one-shot and repeated games.

Figure III.5 below presents the dynamics of average contributions. The difference between treatments is clear: average contributions in L-HL groups remain high over the whole course of the repeated game in Exo treatments but fall rapidly in the Endo treatment (although average contributions started at nearly the same level).

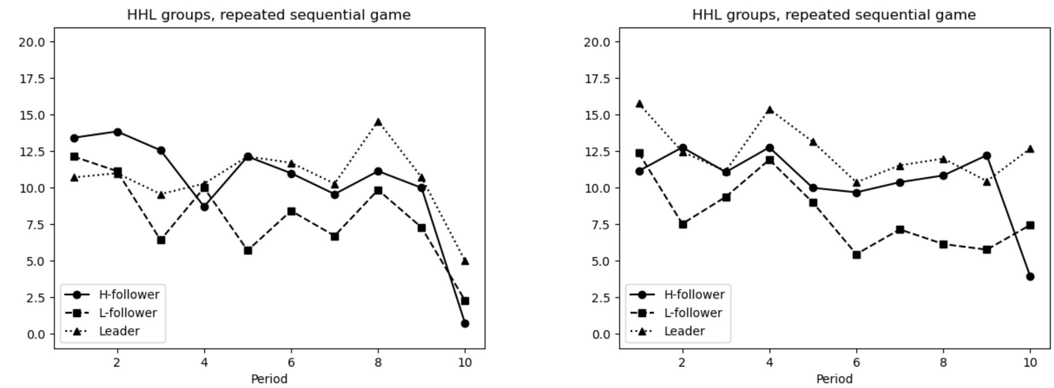
Figure III.6 below plots the dynamics of average contributions of leaders and followers. According to the graph, poor cooperation in L-HL groups in the Endo treatment is the result of no reciprocation for the leader's contribution among followers: although L-leaders started with full contributions, followers did not contribute much, so overall contributions fell immediately.

[16] The coefficients for the dummy indicating that the follower voted for the leader to get a high return have large magnitudes but low statistical significance among H-followers and both high magnitudes and high statistical significance among L-followers. Nonetheless, it is unclear whether followers' votes cause them to contribute differently, or if voting decisions are driven by particular contribution preferences and serve as a signal for the preferred leader's return.

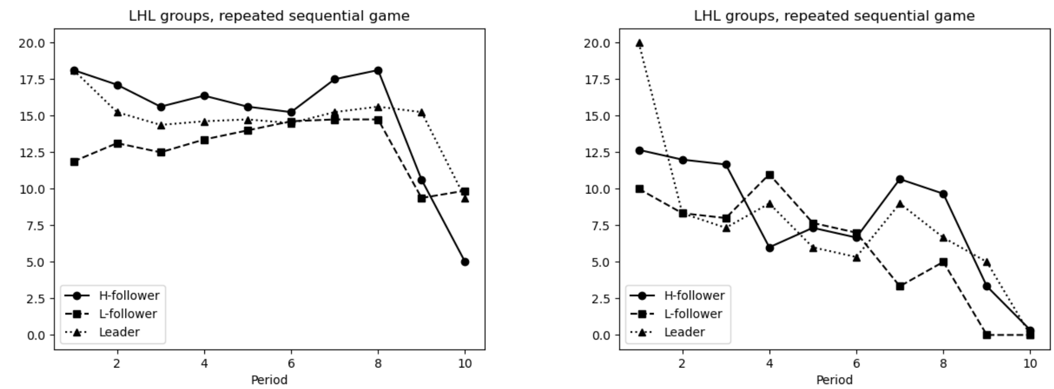


(a) Exo treatments (b) Endo treatment

Figure III.5: The dynamics of average contributions between Exo and Endo treatments



(a) H-HL groups, Exo treatments (b) H-HL groups, Endo treatment



(c) L-HL groups, Exo treatments (d) L-HL groups, Endo treatment

Figure III.6: Comparison of dynamics of average contributions in repeated games

III.7 Concluding Discussion

In this paper, I compare contribution decisions in a sequential public goods game between groups that differ in the composition of returns. I study how the effectiveness of such an institution as leading-by-example changes when the composition of returns among group members is endogenized. I contrast the results obtained from the two experiments that differ in the way in which leaders' returns are established. In one experiment, leaders have predetermined returns (Exo treatments), while in another experiment followers decide upon the return of the leader in their group by voting (the Endo treatment).

My findings suggest that both the composition of returns in the group and the possibility to influence this composition play an important role in contribution decisions of all players. In groups that have predetermined leaders' returns, followers are more willing to cooperate with leaders who benefit relatively less from the public good. There is no such a clear pattern in groups with endogenous leaders' returns: the possibility to choose leaders' benefit has a debatable impact on overall cooperation. On the one hand, followers expect leaders with higher returns to contribute more to the public good, reciprocating followers' voting decisions. On the other hand, these expectations may create additional incentives for followers to free-ride.

These incentives arise from the relation of contribution and own voting decisions. In the experiment, followers condition their contribution decisions on voting outcomes (i.e. on the selected leader's return). According to Figure III.1, in some cases, followers are able to elicit the vote of the other follower (e.g. if follower 1 voted L but received an H-leader). In these cases, followers might have different contribution preferences than in cases when they received the leader for whom they voted. If votes serve as a signal for the preferred leader's return, followers' contribution responses to contributions of the leader with a particular return should depend only on the return of the leader, and not on their own voting decisions. If contribution decisions vary with different voting decision, one may expect that votes determine contributions. I provide evidence in favor of the latter hypothesis: I find that L-followers who *did not* vote for their leaders to have high returns follow their leaders quite closely, but L-followers who did vote for H-leaders seem to ignore leaders' contributions at all. At the same time, H-followers seem to cooperate quite well regardless of their voting decision on the leaders' returns from the public good.

My results suggest that leading-by-example may have spillover effects on the provision of public goods when the composition of returns is endogenized. The possibility to "reward" leaders by other means than responding to their contributions can backfire the intended role model effect, if it allows followers to shift responsibility to the leader. In my experiment, this could potentially happen due to strategic voting by followers: they could vote for a leader to have higher return from the public good to motivate the leader to contribute a larger amount, and lower their own contributions. This is different from the findings from the treatments with the exogenous composition of returns, where the followers are more likely to follow low-benefiting leaders, apparently because their contribution clearly signals the leader's preferred contribution rule and willingness to put their self-interests at risk to guide their groups towards a desired collective goal.

My findings may be particularly helpful for developing behavioral theories of other-regarding preferences. While contribution behavior in homogeneous groups is usually well explained by models of inequality aversion, the differences in followers' reactions to contributions of the leaders observed in my experiments are not. On the one hand, these results are consistent with the idea that followers try to shift responsibility for the provision of the public good to the leader if they can. It seems to be a sort of an excuse "I voted high, so I have done enough" that drives contribution decisions of low-benefiting followers in the treatment with the elected leader's return. In groups where the returns are predetermined there is no such justification, and simple comparison to the player with the greatest (dis)advantage (L- or H-leader or L- or H-follower) seems to dominate other motives. Speaking of justification, the literature on this topic has shown that the weights people assign to fairness considerations differ when the earnings used as endowments in a public goods game are earned or windfall (see Cherry et al., 2005; Kroll et al., 2007). Perhaps, the justification of the way in which the leader's return is established drives the differences in contribution patterns between Exo and Endo treatments in my experiment. The literature on moral licensing (Monin and Miller, 2001) and moral wiggle room (Dana et al., 2007) predicts that players act more self-interestedly in the presence of "excuses". Although I find that shifting responsibility is not punished by the leader, and that the leader reciprocates high returns by increased contributions (cf. van der Weele et al., 2014), in practice leaders can easily anticipate such "excuses" and do not contribute.

On the other hand, such a difference in followers' responses could appear due to the fact that in treatments with exogenous composition of returns I used a between-subject design, while in the treatment with the elected leader's return I used a within-subject design. This can be naturally viewed as a changing reference: other players in between-subject designs versus the player himself in a different state of the world in within-subject designs. Perhaps the most promising direction for further theoretical investigation of cooperation behavior in asymmetric social dilemmas is through models that incorporate an underlying reasoning for both the choice of a specific peer (like in Kießling et al., 2018) and contribution/performance reference (see Falk and Knell, 2004).

Chapter IV

The Bonding Effect of Deferred Compensation — Worker Separations from a Large Firm in Early Transition Russia*

Joint with Thomas Dohmen and Hartmut Lehmann

IV.1 Introduction

In deferred compensation schemes, workers are paid part of their compensation for contemporaneous production in later periods. Often, deferred payments come in the form of rising tenure-wage profiles. These schemes create incentives for workers to continue the employment relationship to capture these rewards. Several theoretical models have shown that deferred compensation schemes discourage turnover and encourage investment and effort provision (see Becker and Stigler, 1974; Lazear, 1979, for early accounts on deferred compensations).[1] In practice, it is typically difficult to write explicit and enforceable contracts on deferred compensations.[2] Workers will therefore only accept such compensation schemes if they believe that the firm sticks to its promise of future payments. Since in the absence of legally binding contracts, firms have an incentive to renege on wage premia of older workers, the literature suggests that reputation serves as an effective com-

*We are grateful to Vladimir Gimpelson, Rostislav Kapeliushnikov and Aleksey Oshchepkov and seminar participants at IZA and at HSE University for valuable comments and suggestions. The usual caveat applies. Dohmen gratefully acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project A01) and Germany's Excellence Strategy - EXC 2126/1- 390838866. Lehmann is grateful to the University of Bologna for granting a sabbatical in Moscow during which this paper took shape. He also acknowledges the financial support provided by the Russian Academic Excellence Project '5-100' within the framework of the HSE University Basic Research Program.

[1] The idea of rising wage profiles in workers' finite careers dates back to Becker and Stigler (1974), who proposed a bonding scheme in which the firm awards workers wage premia throughout their terminable careers until they are detected shirking. To compensate for the discounted value of wage premia, the firm charges an entrance fee of equal value. Lazear (1979) continued this line of research and showed that countless increasing wage profiles exist that satisfy the essential feature of deferred compensation schemes, namely that the discounted value of productivity equals the discounted value of wages.

[2] For example, deferred compensation schemes that promise rewards if performance is adequate are often not enforceable in court.

mitment device (see Akerlof and Katz, 1989; Carmichael, 1989; MacLeod and Malcomson, 1989, 1993, 1998). Despite advances in theory there is scant empirical evidence both on the bonding effect of deferred compensation schemes and on the role of reputation.[3]

In this paper, we provide such evidence using unique personnel data of a large Russian manufacturing firm that include wages, wages withheld (wage arrears) and the exact timing of workers' entry into and exit from the firm. We are thus in a position to study bonding effects of deferred compensation schemes using wage arrears as a special type of deferred payment. Wage arrears occurred in Russia and in this firm during the early 1990s when inflation was high. Whenever our firm withheld wages, it repaid them after some months in their entirety, but it did so in nominal terms. In a high inflation environment this meant, of course, large real losses to all affected workers, presumably leading to a loss in reputation for this firm. Having individual data on workers' separations, we thus can also study how workers who experienced deferred wage payments respond to a deteriorating reputation of the firm.

Previous research on wage arrears has hardly touched on the deferred payment aspect. Naturally, this is unsurprising since researchers wanted to understand why the wage arrears experienced by a majority of workers in CIS countries in the 1990s existed and persisted for long periods. As data were available above all for Russia, we briefly report on the most pertinent studies on this practice in Russia. Layard and Richter (1995a,b) see wage arrears in Russia as a tool to trade-off wages for continuous employment, while Lehmann et al. (1999) and Desai and Idson (2000) perceive arrears as an adjustment mechanism to cushion negative demand shocks or accompany firm restructuring when mass layoffs are considered undesired by the central government. Lehmann et al. (1999) also investigate how local labor market conditions affect the quitting behavior of workers hit by wage arrears, showing that workers who experience wage arrears only quit firms in regions where there are many outside opportunities. The interaction of wage arrears and local labor market conditions in Russia is addressed in the studies of Earle and Sabirianova (2002), and Earle and Sabirianova Peter (2009). Employing matched employer-employee data, they establish that workers' quitting behavior is inversely related to the extent of the practice of wage arrears in the local labor market. In a companion paper, Earle et al. (2010) lucidly discuss why the massive practice of wage arrears in Russia in 1991–1998 can be considered a successful experiment in the normalization of deviant organizational practices.[4] It is worthwhile stressing that the cited studies all assume and in some cases show that wage arrears are deferred payments, i.e. that the owed wages are eventually paid out to workers.

[3] The literature provided indirect evidence on whether firms offer pensions that are designed to retain and motivate workers (Wise, 1985); and whether mergers or high bankruptcy probabilities are associated with flatter wage profiles (Gokhale et al., 1995). Furthermore, Huck et al. (2011) show in a lab experiment that "workers" provide more effort when the "employer" has a history of making generous wage offers in later periods of an employment relationship, but not in response to past honesty.

[4] Our firm practiced wage arrears only as long as it had severe liquidity problems, it did not engage in deviant practices because other firms did so. Therefore, the often criticized "neoclassical" explanation of wage arrears as a temporary adjustment mechanism to large negative demand shocks told for example by Lehmann et al. (1999) and Desai and Idson (2000) seems applicable here.

It is also noteworthy that only the study by Earle and Sabirianova (2002) briefly discusses bonding effects of wage arrears and altered tenure-wage profiles. However, the authors do not directly test bonding effects since their data do not permit such a test.

Wage arrears are a special form of deferred payments, as workers are promised to be paid part of their wages at a later time. Akerlof and Katz (1989) prove that implicit contracts with deferred payments do provide incentives for workers to remain, but require an assumption that firms must not deprive workers of the promised wages. This assumption has been addressed in the theoretical literature, showing that for implicit contracts with deferred compensations to survive, firms optimally decide to pay the promised wage to maintain their reputation (see Bull, 1987). Moreover, MacLeod and Malcomson (1989) demonstrate that, as implicit contracts, deferred payments crucially depend on workers' beliefs about how the firm executes these implicit contracts.

As long as workers believe that withheld amounts are eventually paid out by the firm, wage arrears are expected to bond workers to the firm if workers would forego the promised repayment upon leaving. When wage arrears occur for the first time and workers have no reason to believe that the firm will not repay them, the likelihood of a worker to separate should fall.[5] However, once workers learn that repayments fall short of their expectations, workers feel betrayed. The firm loses its reputation, and consequently the bonding effect of wage arrears fades away. Workers could then perceive the onset of delayed payments of wages as a threat of real income losses, and hence as a sign that it is better to leave the firm before being affected by wage arrears. The separation rate of hitherto-unaffected workers would therefore rise when the firm reverts to wage arrears again. However, we note that the weakening of the bonding effect does not necessarily imply that the bonding effect of deferred compensation completely disappears. This is because workers whose wages are withheld might still be better off waiting to receive at least some money back, as the alternative is to lose the whole withheld amount. The separation decision of workers affected by wage arrears in an environment in which they do not expect full repayment depends on outside options, the expected real income loss, which is determined by the timing and extent of accumulated future arrears and repayments.

These theoretical predictions are borne out by the personnel data of our firm, for which we observe main episodes of wage arrears in 1992 to 1995. Our analysis reveals that wage arrears have a very strong bonding effect on workers during the first episode.

[5] The timing of the first wage arrears cycle in our firm coincided with a severe crisis of the main enterprises in the local labor market. These enterprises like our firm were connected to the military industrial complex. The reform government at the beginning of 1992 tried to address large macroeconomic disequilibria in the Russian economy through price liberalization and by slashing government subsidies to the military industrial complex. This led to a collapse of the demand for output produced by firms connected to the military industrial complex in Russia and our local labor market. Firms then used wage arrears as one tool to adjust to this collapse in demand. Since the macroeconomic reform measures took some time to affect the real economy and since we observe the appearance of wage arrears for the first time in April 1992 in our firm, it is highly unlikely that at this time other firms operating in the same local labor market had already gone through a cycle of wage arrears and repayments. Hence, it is unlikely that workers in our firm had learned from workers in other firms about potential losses in real terms due the repayment of deferred wages in nominal terms.

In fact, no worker whose wages are withheld separates in this first period of wage arrears. When repayments are made in nominal terms, workers incur large losses in real terms and learn that wage arrears are deferred wage payments that are only partially repaid in real terms. This weakens the bonding effect of wage arrears, as becomes visible when wage arrears reappear in a second episode shortly after the repayments of the first wage arrears are completed. Many workers separate preventively at the beginning of the second episode, when they observe other workers having wage arrears, while not yet affected themselves. While some workers whose wages were withheld in the second episode now separate even before the firm repays them, the bonding effect does not disappear entirely. The separation rate of affected workers is still lower, arguably because some workers consider themselves better off remaining with the firm since immediate separation would for all practical purposes imply the complete loss of the withheld wages. Importantly, the separation rate of workers rises sharply once no further repayments are expected.[6]

The rest of the paper is organized as follows. Section IV.2 describes the firm and the personnel data in some detail honing in especially on the wage arrears data. In Section IV.3, we briefly discuss our empirical strategy and report our findings. The same section also discusses the results from robustness checks and alternative mechanisms that could drive our results, while Section IV.4 concludes.

IV.2 Context and Data

We analyze unique data from the personnel records of a large Russian firm, covering 1990 to 2006. The firm is located in a provincial city not far from Moscow and operates in the sector "machine building and metal works". The firm was founded in the 1950s and was in Soviet times part of the military industrial complex, producing military hardware. It was privatized in 1992 using "insider privatization", i.e. giving shares to managers and workers who worked in the firm at the time of privatization. Like in many Russian enterprises ownership of shares became rapidly concentrated in the hand of a few managers in our firm. Already in 1992 when the first period of wage arrears occurred the CEO and a few top managers had a majority of shares under their control. Even though there was collective bargaining on paper in the firm, trade union representatives had virtually no influence on wage policy from early on in the transition, and wage determination was entirely the domain of top management.

During 1992 the reform program of the Gaidar government started to have a strong impact above all on firms connected to the military industrial complex. To combat large

[6] The Russian labor code in the 1990s already foresaw as it does now that firms have to pay their workers and not paying them was and is illegal. However, unlike today, in the 1990s there were no legal tools to enforce the law that declared wage arrears an illegal practice. Consequently, there were very few cases where workers went to court to assert their claims of unpaid wages. Even in the case of a favorable judgment, given the length of the court procedure and given the high inflation environment workers only received a small fraction of the owed wages in real terms. Therefore, when in the first half of the 1990s a worker leaves a firm that owes her or him wages, she or he forgoes these owed earnings for all practical purposes.

macroeconomic disequilibria in the Russian economy, the government liberalized prices of many goods and services and slashed the budget of the military and subsidies to enterprises. Hence the demand for military hardware collapsed and firms' losses were no longer compensated with subsidies coming from the central state budget. Consequently, like many other firms in the military sector our firm had to convert its production to civilian goods, which of course could not be done overnight. The CEO and top management decided to convert the firm's production to well equipment for gas and oil production and smith-press equipment. This conversion process, while successfully accomplished over a period of three years,[7] was painful for management and the workforce leading to severe liquidity problems between 1992 and 1995. As mass layoffs were not an option for enterprises in Russia at the onset of transition (see, e.g. Gimpelson and Kapeliushnikov, 2013), like most Russian firms our firm withheld wages of a majority of workers for extended periods or put some workers on prolonged unpaid leave. Once the firm experienced a steady stream of orders for its new products at the beginning of 1996, such unorthodox methods of adjustment to negative demand shocks disappeared from the toolbox of top management. In actual fact, after 1995 we do not have records of wage arrears in this firm.[8]

In the early years of transition, the Russian labor market was in major turmoil, characterized by excessive turnover, caused predominantly by very large quit rates of workers. Many of them had the perception that better earnings opportunities existed outside the enterprises where they had worked thus far (Lehmann and Wadsworth, 2000). Table IV.A.1 in the appendix shows for our firm the composition of the workforce across six categories,[9] the evolution of the total number of employees and average real monthly wages for the period 1991 to 2006. In column 8 of the table, we observe the tremendous reduction of employees (roughly one-third of the original workforce) that occurred between December 1991 and December 1995 when the firm had finished the conversion process into civilian goods production. This net change in employment is driven by a separation rate of 75 percent and a hiring rate of 34 percent over the period. Virtually all separations were voluntary quits which implies that many workers in our firm must have seen better earnings opportunities or work conditions outside the firm. At the same time, the firm hired many new workers, replacing those who had left the firm[10] and must, therefore, have been aware of the importance of its reputation regarding the fair treatment of its workers.

The data record net monthly wages, wage arrears, and repayments of wage arrears. Separately recorded are monthly and annual bonuses and monthly working hours. Net

[7] For this successful conversion to civilian production, the CEO of the firm was ranked among the top 35 managers in the Russian machine-building industry by Russian business magazine Kommersant (2006).

[8] However, for the Russian economy and industry at large wage arrears still grew after 1995 and reached a peak in 1998 during the financial crisis of that year, after which they were rapidly reduced (Gimpelson and Kapeliushnikov, 2013).

[9] In Soviet industry we find the six categories of workers given in the table. Even today, this classification is still used in many enterprises. It is also noteworthy that the composition of the workforce remained roughly the same over the entire reported period.

[10] Since the educational composition of those workers who left the firm in 1992 to 1995 is very similar to the educational composition of the workers hired in this period we are confident that the firm engaged above all in replacement hirings.

monthly wages include all additional payments made in a particular month, like monthly bonuses for earlier completion of the planned work or annual premia. Data on wage arrears and repayments are given for the early transition period from 1992 to 1995, when wage arrears were a problem in the firm, while the data on wages and hours worked are available from 1990 to 2006. A rich set of demographic characteristics includes gender, birth date of all workers, marital status, number of workers' children and their birth dates, as well as workers' educational attainment. Finally, for those workers who entered or left the firm we have the exact dates when this occurred. We can thus very precisely link the occurrence of wage arrears and their repayments with the occurrence of workers' separations.[11]

Let us conclude this section with a thorough discussion of the wage arrears data that we have at our disposal. We have exact information on the amounts of withheld wages and on the months when this happened as well as on the amounts of repaid wages and on the months when these repayments occurred.[12] Figure IV.A.1 in the appendix shows the dynamics of monthly inflation starting one year before the reform policies were implemented in Russia (January 1992). By the standard definition of hyperinflation, i.e. the monthly inflation rate is greater than 50 per cent, we see three episodes: March 1991, January and March 1992. Wage arrears started in April 1992 as inspection of Table IV.A.2 in the appendix demonstrates after the last hyperinflation episode. However, from Figure IV.A.1 we can also infer that from April 1992 until the end 1994 when wage arrears and their repayments mainly took place in our firm monthly inflation oscillated between 10 and 40 percent. Therefore, obviously nominal wages were withheld and repaid in a high inflation environment.

The upper panel of Figure IV.1 presents the aggregate amounts of wage arrears and their repayments in nominal terms, while the lower panel shows these amounts in real terms. We see two major episodes of wage arrears, from April 1992 to December 1992 and from August 1993 until December 1993, while repayments are from January to March 1993 following the first episode, and evenly spread out over the whole year of 1994 following the second episode. There is a third episode of wage arrears in 1995, but as the lower panel makes clear these arrears are negligible in real terms. The firm repays all wages that it withholds in nominal terms; this is very clear from the upper panel of Figure IV.1 since the area of black bars (wage arrears) is identical to the area of gray bars (repayments) in

[11] Our records also contain detailed descriptions of all positions each worker ever held in this firm, except for the top management jobs. The records include position title and code, as well as the department and the subdepartment to which the position belongs. Such detailed descriptions allow us to identify all position changes in the firm, and thus to account for internal mobility.

[12] The above cited studies on wage arrears that inter alia also focus on the relationship between wage arrears and workers' quitting behavior do not have such precise data at their disposal as we have in this study. While we have the exact cumulative amount of wage arrears and wage repayments for each worker, the studies based on the Russian Longitudinal Monitoring Survey (RLMS) use accumulated owed months of wages and accumulated paid off months of wages, which can be considered good proxies for accumulated wage arrears and repayments. More worrisome is the measure that captures workers' quits. Since employment state and tenure in the RLMS are only given annually in the reference week, it is not possible to exactly determine the date of workers' separations like we can with our personnel data. Therefore, measurement error is a concern in these studies.

both episodes.[13] However, the situation is very different in real terms, as inspection of the lower panel of Figure IV.1 shows. The firm pays back only a fraction of the withheld wages: in fact, in the two episodes these fractions are 54 per cent and 48 percent, respectively. While in absolute terms wage arrears are larger in episode 1, the relative burden is heavier in the second episode.[14] Of course, how the evolution of wage arrears and their repayments affects workers' behavior is the main focus of our study.

IV.3 Analysis

IV.3.1 Descriptive Analysis

The prevalence of wage arrears during the first two episodes was massive, as more than 90 percent of the workforce was affected. Table IV.1 shows the share of individuals by employment category whose wages were withheld at the end of particular periods. For illustrative purposes we split the first episode into two sub-periods because the timing of wage withholdings differed strongly across employment categories during the first episode (see Table IV.A.2 in the appendix for summary statistics on the incidence of wage withholdings by month). In particular, the large majority of accountants, engineers, and auxiliary production workers had wage arrears from April on, while other employment categories were affected not before October 1992.[15] In the first episode, the majority of the workforce were repaid the withheld wages in January 1993 (see Table IV.A.3 in the appendix for summary statistics on the incidence of completed repayments by month).[16] Nevertheless, payments of outstanding wages to service workers finished only in March 1993, and some workers were not fully repaid until June 1993.

The second episode starts shortly after the last repayments have been completed in the first episode: wage withholdings begin in August 1993 and occur until December 1993, while the repayment period spans the entire calendar year of 1994. In the second episode, the incidence of wage withholdings was quite uniform, as the majority of the workforce (more than 80 percent) was affected from the beginning of the episode (see also Table IV.A.2 in the appendix).[17] Most workers only received their nominal wages back in full a year later (see Table IV.A.3 in the appendix). The third episode of wage

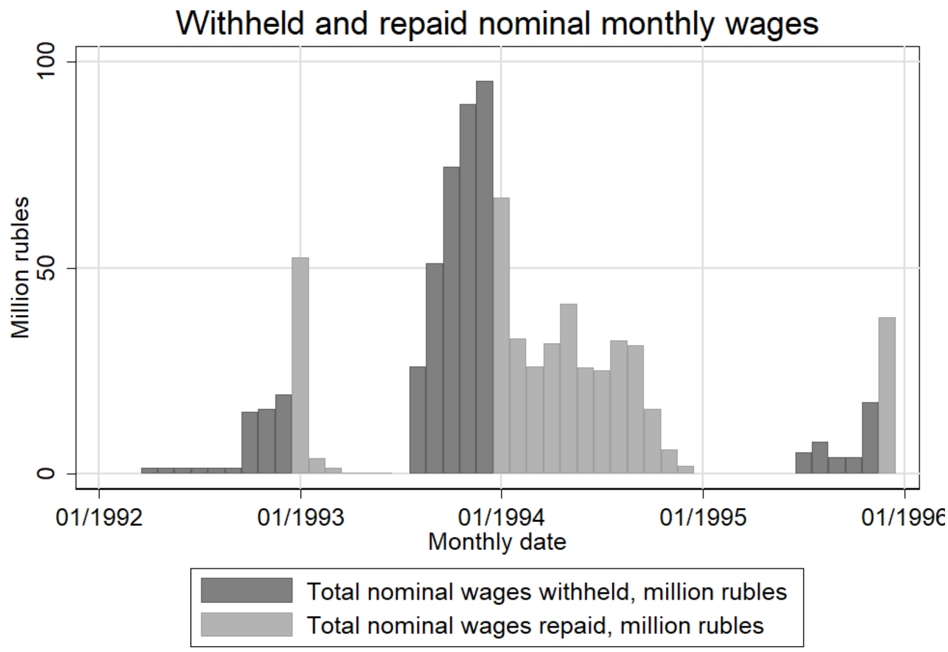
[13] In episode one the firm withholds 57.58 million rubles and repays 57.56 million rubles, while in episode two the respective numbers are 336.89 and 336.23 million rubles. In the second episode the numbers diverge somewhat because 251 workers leave the firm before they are paid the wage arrears in their entirety.

[14] Wage arrears in absolute terms are much larger in episode 1 because real wages are double to what they are during episode 2, as inspection of the last column of Table A.1 shows.

[15] According to Table IV.A.2, by December 1992, the last month when wages were withheld in the first episode, the firm owed wages to more than 90% of its workers. In September 1992, only 11% of managers and 9% of primary production and service workers were affected.

[16] According to Table IV.A.3, 90% of managers, engineers and production workers received their nominal wages back in January 1993, but only 70% of accountants did. At the same time, nearly 95% of service workers only received the withheld amounts in March 1993.

[17] Table IV.A.2 shows that in the second episode, 80% of managers and accountants received wage arrears in August 1993, but only about two-thirds of other workers were affected. However, in December 1993, the share of workers whose wages were withheld increased to more than 95%.



(a) In nominal terms

(b) In real terms (Jan 1992 rubles)

Figure IV.1: Monthly aggregate amounts withheld and repaid

arrears occurs in the second half of 1995, with wage withholdings taking place from July to November, and repayments all done in December 1995. Notably, only a small share (less than 10 percent) of the workforce was confronted with wage arrears in the third episode, and managers were not affected at all.

Table IV.1: Average share of workers whose wages were withheld in corresponding periods, %

Category	Episode 1 Apr'92–Sep'92	Episode 1 Oct'92–Dec'92	Episode 2 Aug'93–Dec'93	Episode 3 Jul'95–Nov'95
Accountants	77	96	98	5.9
Managers	11	97	97	0
Engineers	83	94	94	11
Primary production workers	8.9	96	97	10
Auxiliary production workers	82	93	95	9.7
Service staff	13	93	95	13

Figure IV.2 depicts periods of wage withholdings and wage repayments. Dark gray bars represent the monthly share of the workforce whose wages are withheld in a given month, while light gray bars represent the share of workers who receive some repayments in a given month. Finally, empty contoured bars represent the share of workers who are fully repaid in a given month. The figure also plots monthly separation rates as a line (scale on the right axis).[18] The thick black line represents the overall monthly separations. Thin dashed lines correspond with separations among workers who were affected by wage arrears in the corresponding episode, and among those workers who were not, respectively.

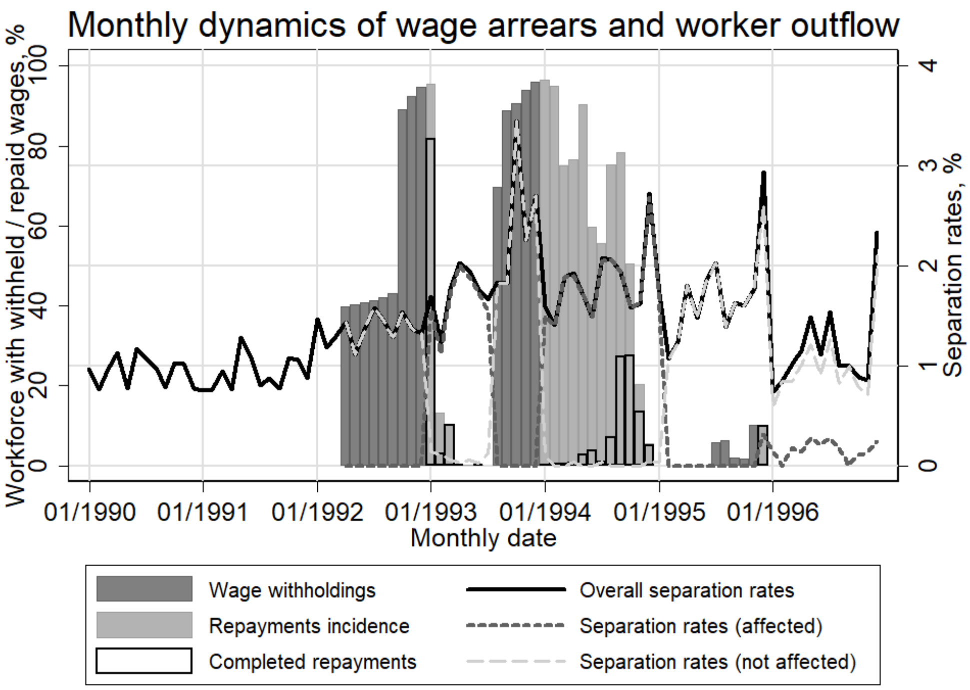


Figure IV.2: Wage withholdings, their repayments, and worker separations

[18]We calculated monthly separation rates as the number of workers who left the firm in a given month divided by the total number of employees in that month. The reasons for job leaves are not reported in our data.

There are several notable observations about the timing of wage arrears and separations to be made from this graph. First, during the first episode lasting from April 1992 to June 1993, overall monthly separations remained stable after an increase in January 1992, when the liberalization of prices began. However, note that separation rates among workers affected by wage arrears (the short-dashed gray line) appear to be equal to zero up until January 1993, when the majority of the workforce receive repayments. Table IV.2 below provides more support to this observation: during the period of wage withholdings in the first episode, only workers not affected by wage arrears separate from the firm. Since no worker affected by wage arrears separated during the period of wage withholdings in the first episode, it appears that wage arrears had a strong bonding effect.

Table IV.2: Worker separations by the incidence of wage arrears

	Not Affected		Affected	
	per month	total	per month	total
Period				
Episode 1, withholdings	53	473	0	0
Episode 1, repayments, not yet repaid	2.5	15	.33	2
Episode 1, repayments, repaid fully	0	0	58	348
Episode 2, withholdings	78	389	0	0
Episode 2, repayments, not yet repaid	.5	6	21	251
Episode 2, repayments, repaid fully	0	0	34	409
Episode 3, withholdings	49	243	0	0
Episode 3, repayments, not yet repaid	75	75	0	0
Episode 3, repayments, repaid fully	0	0	9	9

Notes. Reported are average monthly and total numbers of workers separating during the corresponding period, divided into two groups: those who were affected anytime during the corresponding episode of wage arrears, and those who were not.

This result is well aligned with theoretical predictions about the bonding effect of deferred payments. At the beginning of the first episode, workers have no reason to believe that the firm will not repay them.[19] It, therefore, seems that the increase in separation rates during the period of price liberalization might appear slower in our firm compared with the local labor market due to the bonding effect of wage arrears.

Second, it is notable that the overall separation rate markedly increases (by 0.59 percentage points in monthly separations, i.e. roughly a 40 percent increase) towards the end of the first episode, namely after the firm repaid wage arrears to most of its workers. The rise in separation rates might reflect that affected workers are more likely to quit when no further repayments can be expected, either because they make up for intended separation that they postponed until they were repaid, or because they lose motivation as they perceive the firm's repayment policy as unfair in the wake of high inflation. After

[19] We do not observe wage arrears in 1990–1991 (the two earliest years in our personnel data set). Even if there were wage arrears before, inflation was negligible due to controlled prices, so payments in nominal terms and real terms coincide.

workers receive the withheld wages in the first episode, they realize that wage arrears do cause real income losses, as workers are not repaid fully in real terms due to high inflation.[20]

However, in our analysis below, it will become apparent that increasing separation rates are not driven by the quitting behavior of affected workers, but rather they mainly stem from employees who separate preemptively, potentially fearing another period of wage arrears. The perception that wage arrears reduce real wages becomes particularly relevant when wage arrears reappear: in the second episode of wage arrears (starting from August 1993), we see that overall separations jump up right at the beginning of the new episode.[21] The bonding effect of wage arrears fades because hitherto non-affected workers start fearing real income losses. Table IV.2 also suggests that many workers affected by wage arrears in the second episode separated even *before* they are fully repaid the withheld wages in nominal terms, providing another piece of evidence in favor of the weakening of the bonding effect.

IV.3.2 Empirical Framework

Next, we scrutinize whether the observed separation patterns in the raw data are robust to controlling for potentially confounding factors such as individual characteristics or time-specific effects. Since we are interested in the determinants and the timing of job separations, we estimate the hazard rate of job separation (i.e. that a separation occurs at time t given that the worker who entered the firm at time t_0 had been employed up to time t) conditional on a set of observable characteristics X. The hazard rate of job separation is given by

$$\lambda \equiv \lambda(t, t_0, X) = \lim_{\varepsilon \to 0} \Pr\{t_1 \in [t, t+\varepsilon] | t_1 \geq t, t_0, X\}/\varepsilon, \tag{IV.1}$$

Since the economic environment in which the firm operates is non-stationary (i.e. marked by political and (macro)economic shocks) calendar time dependence of job duration, and hence of separation rates, is to be expected. In a non-stationary environment, exit rates are likely correlated because turnover might be governed by firm policy or by shocks that affect workers in similar ways at a given calendar time. By reversing the role of calendar time and employment duration in a Cox proportional hazard model (Cox, 1972, 1975), we can account for a correlation between exit rates at the same (calendar) time for different individuals (proposed by Imbens, 1994). Such a model can be specified as

$$\lambda(t, t_0, X) = \lambda_0(t) * f(z(t - t_0), X; \theta), \tag{IV.2}$$

[20] Yearly inflation was 440.6% in 1992 (with a peak of 170% in January 1992), 252.8% in 1993, and 118.1% in 1994; see also Figure IV.A.1 in the appendix for the dynamics of monthly inflation rates. Note that these figures correspond to local (regional) CPI, assessed in 1997. Later, the Russian Statistical Agency has changed its methodology to calculate inflation rates, so current reports on monthly inflation for that period are lower.

[21] Mann-Whitney U-test rejects the hypothesis that monthly separation rates during wage withholdings are equal between the second episode and other episodes ($p < 0.01$).

where θ is a vector of parameters to be estimated. All events that affect separation rates of workers in the same way on a given day and that are not accounted for by a set of characteristics in X are captured by the baseline hazard $\lambda_0(t)$.

We further assume that $f(\cdot)$ is separable in terms of duration dependence and other observable characteristics, and that $f(\cdot)$ is the exponential function. We specify z as

$$z(t-t_0) = \exp \sum_{i=1}^{N} \alpha_i I\left[d_i < t - t_0 \leq d_{i+1}\right], \tag{IV.3}$$

where $I[\cdot]$ is the indicator function, and $d_{i+1} - d_i$ denotes a particular period length.[22]

We therefore estimate the following model:

$$\lambda(t, t_0, X) = \lambda_0(t) * \exp \sum_{i=1}^{N} \alpha_i I\left[d_i < t - t_0 \leq d_{i+1}\right] * \exp\left(X\beta\right). \tag{IV.4}$$

The set of explanatory variables X includes a set of worker characteristics, such as age, gender, educational attainment, marital status, the presence of children in the household, and a set of job characteristics, including dummies for job levels and indicators for intra-firm mobility.[23] Since we are interested in the effects of the firm's wage arrears policy on affected and non-affected workers, we include dummy variables in X that indicate the incidence and timing of wage arrears.

In order to assess differences in the effect of wage arrears on separations between the first and second episode of wage arrears, we estimate the model separately for each of these two episodes. In our Cox proportional hazard regressions for the first period of arrears, we define an indicator variable that takes on the value of 1 once the worker becomes affected by wage arrears and that remains being equal to 1 until nominal arrears are fully repaid. We uncovered in our analysis in Table IV.2 that no worker affected during period 1 separated from the firm before repayment started, but that a few workers separated before being fully repaid. Hence, this dummy variable is identified from those who separate without being fully repaid.

For the second episode, we estimate two Cox proportional hazard models. In the first one, we use the same specification as for the first period. In addition, we estimate a model in which we add a dummy variable that indicates whether a worker had been affected in the first episode. This variable captures whether previous experience with wage arrears affects behavior. We conjecture that those who were affected by wage arrears in the first period have learned from personal experience and therefore have a stronger motive to separate in the second periods when wage arrears re-emerge. The coefficient estimates of

[22] We are not primarily interested in duration dependence, which is captured by $z(t-t_0)$, and therefore only approximate it by a step function in firm tenure.

[23] Age, tenure, and education control variables were transformed into the sets of categorical dummy variables to allow for more flexible model specifications. We divide tenure duration into sixteen categories, age – into 49 categories, education – into four (high school or lower, professional education, higher education, and those who are currently in the process of obtaining the college degree).

these indicator variables are our main variables of interest.

IV.3.3 Estimation Results

We start by assessing how the incidence of wage arrears affects separation rates in both periods. According to our hypothesis and in line with the descriptive analysis above, we expect strong bonding effects (i.e. lower separation rates of affected workers) in the first period and weaker but positive bonding effects in the second period. Table IV.3 shows the estimated coefficients for dummy variables for being affected by wage arrears and for being experienced with wage arrears in the first episode from our first model.

Table IV.3: The effect of having wage arrears

Dependent variable: hazard rate at t	(1) Episode 1 Apr'92–Jul'93	(2) Episode 2 Aug'93–Jan'95	(3) Episode 2 Aug'93–Jan'95
1 if affected at t	-3.824***	-3.703***	-3.710***
	(0.158)	(0.098)	(0.099)
1 if experienced arrears in the first episode			2.682
			(1.457)
Other controls	Yes	Yes	Yes
Observations	58569	55253	55253

Notes. Cox proportional hazard estimates, robust standard errors in parentheses. Other controls (not reported here) include: dummies for education, age, tenure, current position, gender, marital status, and the presence of children, as well as dummies for internal mobility (promotions, demotions, horizontal changes). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The estimates in Table IV.3 reveal that separation rates substantially decline once a worker is affected by wage arrears. In the first period, the hazard rate of job separation among affected workers is 45.8 times lower compared with that of unaffected workers, as the coefficient estimate of -3.824 in column (1) indicates. These results provide strong evidence for the bonding effect of deferred compensation schemes when workers still expect that wage arrears are adequately compensated for in the future.

The estimated effect of other control variables that are not reported in Table IV.3 are consistent with theories of worker turnover (see, e.g. Jovanovic, 1979) and findings in the empirical literature (e.g. Dohmen and Pfann, 2004). We show these estimates in the appendix.[24] We find that younger workers and workers with short tenure at the firm are more likely to separate. Production workers of lower subcategories have a higher estimated hazard rate of job separation than service staff (the reference group), while high-skilled production workers, managers and engineers have lower hazard rates of job separation (see Tables IV.A.7 and IV.A.8 in the appendix). Workers who have been promoted or changed jobs laterally within the firm during the period of 1992–1994 have a smaller separation

[24] Tables IV.A.4, IV.A.5, and IV.A.6 show estimates for tenure and age dummies, respectively. Tables IV.A.7 and IV.A.8 show estimates for firm and career related characteristics, while IV.A.9 documents the estimates for additional individual characteristics.

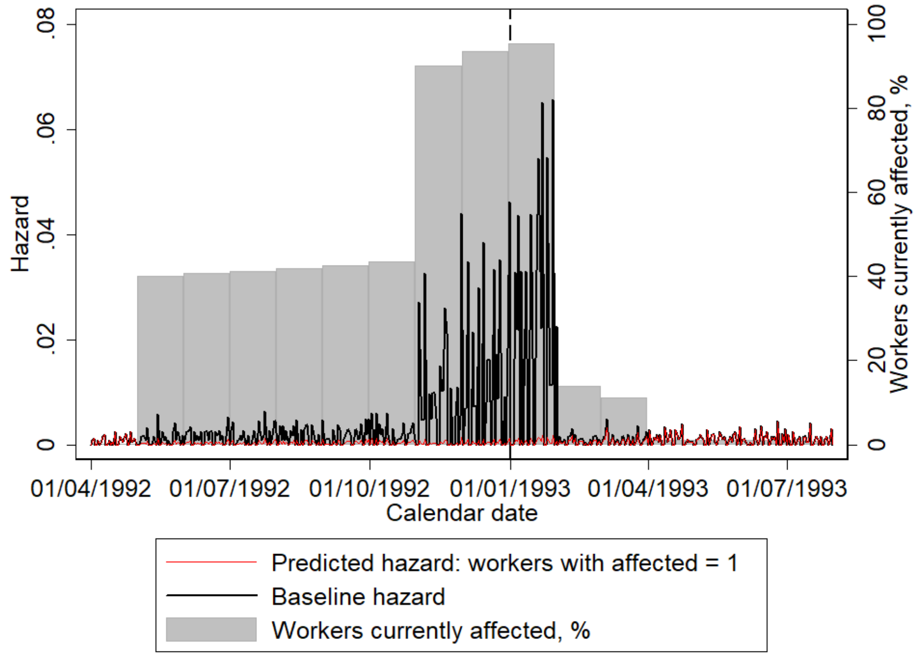
hazard. Highly educated workers tend to separate more likely, possibly, because they have better outside options. Finally, we observe that female workers as well as married workers and those having at least one child are more likely to remain with the firm.

In the second period this relative drop in the job separation hazard is slightly less marked, but the bonding effect remains very strong (see column (2)). On average, affected workers are 40.6 times less likely to separate. Figure IV.3 visualizes this effect by plotting the baseline hazard from the specification of column (1) and column (2) of Table IV.3 along with the proportional (baseline) hazard rates for the workers affected by wage arrears (solid red line) in Panels (a) and (b) respectively. The figure also displays the proportion of workers affected by wage arrears during these two episodes. In addition to what the coefficient estimates in Table IV.3 indicate, the figure suggests that non-affected workers become relatively more likely to separate when wage arrears expand to increasingly more workers. Notably, the baseline hazard increases strongly shortly after the onset of wage arrears in the second period (see panel (b) of Figure IV.3), indicating that workers now perceive wage arrears as disadvantageous and separate preventively to avoid being affected.

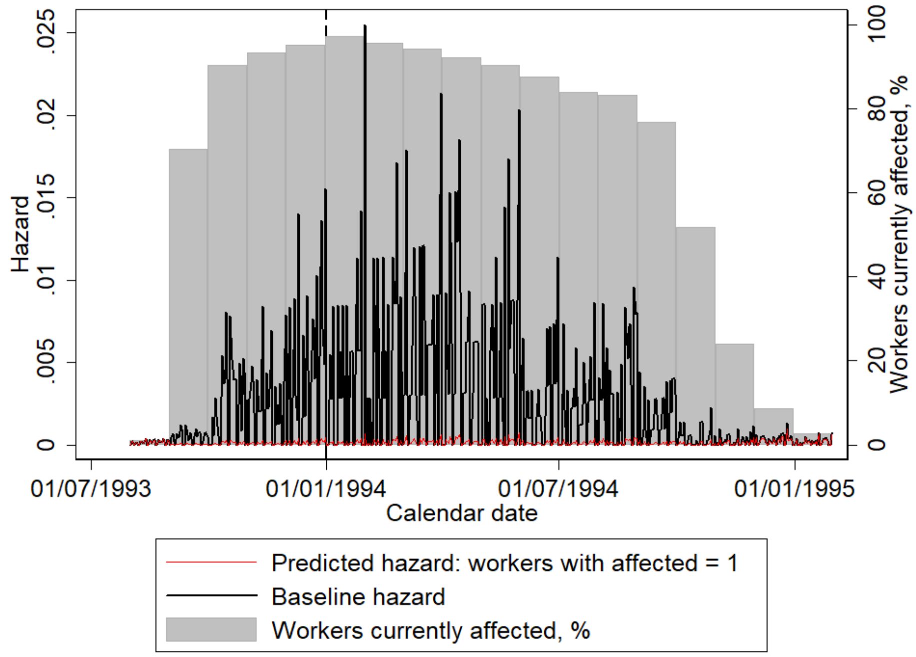
This is a first indication that workers have learned from observing the firm's behavior in the first period so that the bonding effect of wage arrears fades away. The estimates in column (3) of Table IV.3 provide some additional evidence, suggesting that those who had experienced wage arrears themselves in the first period become 14.6 times more likely to separate in the second episode of wage arrears compared with workers who were not subject to wage arrears in the first period. It is important to note that this effect is not statistically significant, likely due to lack of power. However, if we are willing to interpret the very imprecisely-estimated coefficient, we observe a dramatic increase in separation rates of workers who had already experienced wage arrears in the first episode.[25] In sum, our results indicate that deferred compensation backfires when workers fear that it results in income losses. First, when the firm starts using deferred compensation after an episode in which workers could learn that deferred wages cause real income losses, the separation rate rises even among workers that are not subject to wage arrears. Second, the bonding effect of wage arrears fades strongly among workers that previously had experienced that deferred wages result in income losses.

Additional evidence for the fading of the bonding effect comes from an inspection of the baseline hazard rate of a model estimated for the entire period that includes all control variables but the indicator variables for being affected. Figure IV.4 plots this baseline hazard rate and reveals that the risk of separation is higher in the second period of wage arrears than in the first period, indicating that separation rates rise for all workers when the firm starts using wage arrears again after workers had observed that wage arrears

[25] The estimates in column (3) also suggest a weakening in the bonding effect as the hazard rate for the workers who are affected by arrears in the second period and who had already experienced wage arrears in the first period is only 2.8 times lower than the one for workers who were not affected in both periods.



(a) Episode 1 (Jan'92–Jul'93)



(b) Episode 2 (Aug'93–Jan'95)

Figure IV.3: Baseline hazard rates for workers with and without wage arrears estimated from specifications (1) and (2) of Table IV.3

in the first period were associated with real income losses.[26]

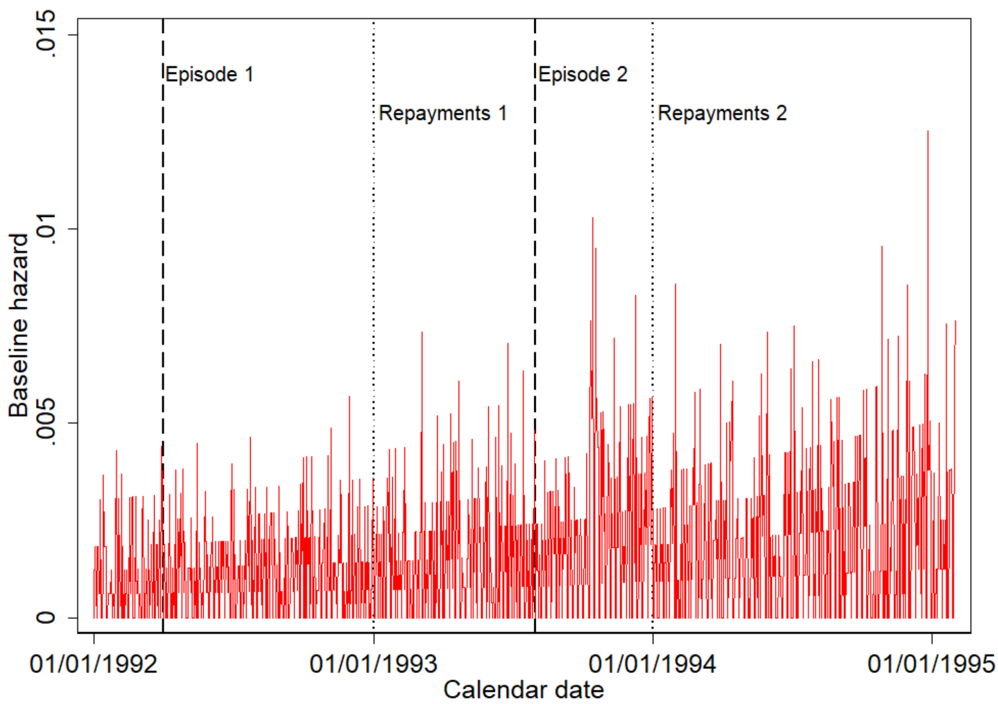


Figure IV.4: Baseline hazard contributions for Jan'92–Jan'95

Notes. Vertical dashed lines show the dates when the first and the second episodes of wage arrears start, correspondingly. Vertical dotted lines show the dates when corresponding repayments in the first and the second episodes of wage arrears begin.

Despite the fact that we control for a large set of observable characteristics, including education, age, gender, marital status, the presence of children, firm tenure, job position in the firm as well as intra-firm mobility, we acknowledge that our estimates might still be biased due to unobserved heterogeneity (e.g. because the incidence and the timing of withholdings or the timing of repayments might depend on unobserved characteristics. However, importantly, since we find very strong effects, even when we condition on a large set of observable characteristics, we are confident that the two main qualitative results hold: (1) deferred wages have strong bonding effects, but (2) these bonding effects fade when firms cannot credibly signal that deferred wages will be adequately compensated for in the future.

In a second set of regressions, we investigate how being repaid affects the separation hazard. For this purpose, we define an indicator variable that switches to one once the worker has been fully repaid. Clearly this variable is not independent of the indicator variable for being affected by wage arrears. Therefore, we split the analysis for each episode of arrears into two sub-periods. The first sub-period comprises the months in

[26] Note that vertical dashed lines indicate the onset of the first and the second episodes of wage arrears. Vertical dotted lines in the figure mark the starting dates of the repayment periods during the first and second episode of wage arrears.

which wage arrears accumulated (i.e. April 1992 until December 1992 and August 1993 until December 1993). As described above, no repayments were made in this period. The second sub-period resembles the repayment period (i.e. January 1993 until July 1993 and January 1994 until January 1995). In the models for the first sub-period, we include an indicator variable for being affected, while in the model for the repayment period we include a dummy variable for being repaid.

Table IV.4 shows the estimated coefficients for these sub-periods. Columns (1) and (3) reveal the strong bonding effect of arrears during each of the sub-periods in which arrears accumulate. In fact, no affected workers separated during this sub-period of the first episode of wage arrears. Columns (2) and (4) indicate that the hazard rate rises dramatically once (affected) workers are repaid their nominal wages, while it does not rise in the first episode. These results corroborate our findings that deferred compensation induces strong bonding effects but that they backfire when they turn out to be disadvantageous for workers, as is reflected in the strong increase of separations at the end of the second episode of wage arrears.

IV.3.4 Additional Results and Robustness

One issue that we have neglected in our main specification above is heterogeneity in the "treatment" effect (cf. Abbring and Van den Berg, 2003). Differences in the severeness of wage arrears, as measured, for example, by the absolute or relative amounts of wages being withheld might be related to differences in the strength of the bonding effect. In order to gauge the relevance of such variation in treatment intensity, we constructed a variable that captures the cumulative percentage of currently withheld wages and a variable that captures the cumulative repayments as a percentage of wages. Moreover, we constructed the difference between the share of withheld wages and the ratio of the repayment to the current wage (accumulated over the two episodes).

Our estimation results in Table IV.5 demonstrate again that wage arrears bond workers to firms in both episodes, the estimated bonding effect of our intensity measure of withheld wages is weaker in the second period, and workers who had experienced wage arrears in the first episode have a higher hazard of rate of separation in the second period.

Table IV.4: The effect of wage arrears in corresponding sub-periods

Dependent variable: hazard rate at t	(1) Episode 1 Apr'92–Dec'92	(2) Episode 1 Jan'93–Jul'93	(3) Episode 2 Aug'93–Dec'93	(4) Episode 2 Jan'94–Jan'95	(5) Episode 2 Aug'93–Dec'93	(6) Episode 2 Jan'94–Jan'95
1 if wages withheld in t	-52.267		-45.518***		-68.129	
	(.)		(0.105)		(.)	
1 if repaid by t		-0.269		3.107***		3.110***
		(0.245)		(0.150)		(0.150)
1 if experienced arrears in the first episode					-20.172	0.954
					(.)	(0.929)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	34220	24349	16153	39100	16153	39100

Notes. Cox proportional hazard estimates, robust standard errors in parentheses. Other controls (not reported here) include: dummies for education, age, tenure, current position, gender, marital status, and the presence of children, as well as dummies for internal mobility (promotions, demotions, horizontal changes). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table IV.5: The effect of the intensity of wage arrears

Dependent variable: hazard rate at t	(1) Episode 1 Apr'92–Jul'93	(2) Episode 2 Aug'93–Jan'95	(3) Episode 2 Aug'93–Jan'95
Cumulative share of withheld wages	-228.675*	-8.901***	-9.587***
	(91.682)	(0.346)	(0.356)
1 if experienced arrears in the first episode			5.596***
			(0.633)
Other controls	Yes	Yes	Yes
Observations	58137	54812	54812

Notes. Cox proportional hazard estimates, robust standard errors in parentheses. Reported are the estimates for the measures of intensity of wage arrears (cumulative share of withheld wages in monthly wages) and the dummy for previous experience with wage arrears. Other controls (not reported here) include the same set of characteristics as described in the notes for Table IV.3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

IV.3.5 Alternative Explanations

One could argue that the firm deliberately jeopardized the bonding effect of wage arrears by only repaying nominal wages to create an instrument that induces voluntary separations when wage arrears are used again. If that was the case we would expect that the firm implements wage arrears for those workers in the first period from whom it wants to separate later.[27] We believe that such a story is not plausible for the following reason. Our estimates suggest that separation rates rise generally during the second episode of wage arrears, i.e. also among workers who were not affected by arrears in the first period. If the firm used its wage arrears policy optimally, this could only be squared with a downsizing motive of the firm. In this case, we would not only expect that the workforce shrinks, but also that separating workers are not replaced. However, this is not what we observe in the data. Figure IV.5 below shows both monthly outflows and inflows, in addition to the shares of workers affected by wage arrears in corresponding months. It becomes apparent that the firm increases hiring after massive separations at the beginning of the second episode: average monthly hiring rates in 1994 are almost two times higher than in 1993 and nearly four times higher than in 1992. We also perform Granger causality test for inflows and outflows and find that outflows Granger cause inflows, but inflows do not Granger cause outflows. Furthermore, we find that repayments Granger cause outflows, but wage withholdings do not. No doubt, Granger causality does not reflect causality in the economic sense but suggests the time sequence of the events, i.e. that outflows precede inflows.

It also seems highly unlikely that the firm used its wage arrears policy purposefully to attach workers. In this case, we would have expected the firm to compensate real earnings losses by other means; for example, through future wage growth or promotions. However, we do not observe this in the data. Real wages only increased once (see Figure IV.A.2 in Appendix IV.A), in January 1993, and have been falling after that, so there was no compensation in the form of increased real wages.[28] Although, we observe that the 1-year growth in log-wages is larger for the workers affected by January 1994 than for non-affected ones, we do not find that affected workers' relative compensation as measured by the rank in the earnings distribution changes after January 1994, when wages arrears came to a halt in the second period (see Appendix IV.B).

Likewise, we do not find evidence that the firm treats affected and non-affected workers differentially when it comes to job promotions. In addition, the absolute frequency of promotions is low relative to the incidence of wage arrears. As a result, the large majority of workers could not expect to be promoted, despite the fact that promotion rates rise during repayment periods. This increase in promotion rates is likely to emerge as separations trigger vacancy chains that are filled by internal promotions and hiring at the entry-job

[27] Let us recall that mass layoffs were not an option for Russian firms in the early years of transition.

[28] According to Figure IV.A.2, average nominal wages increased nearly thirteen times in January 1993 to cope with inflation in 1992. Another substantial increase occurred in January 1994, but real wages never reached the same level as in January 1992. After January 1993, real wages were steadily falling.

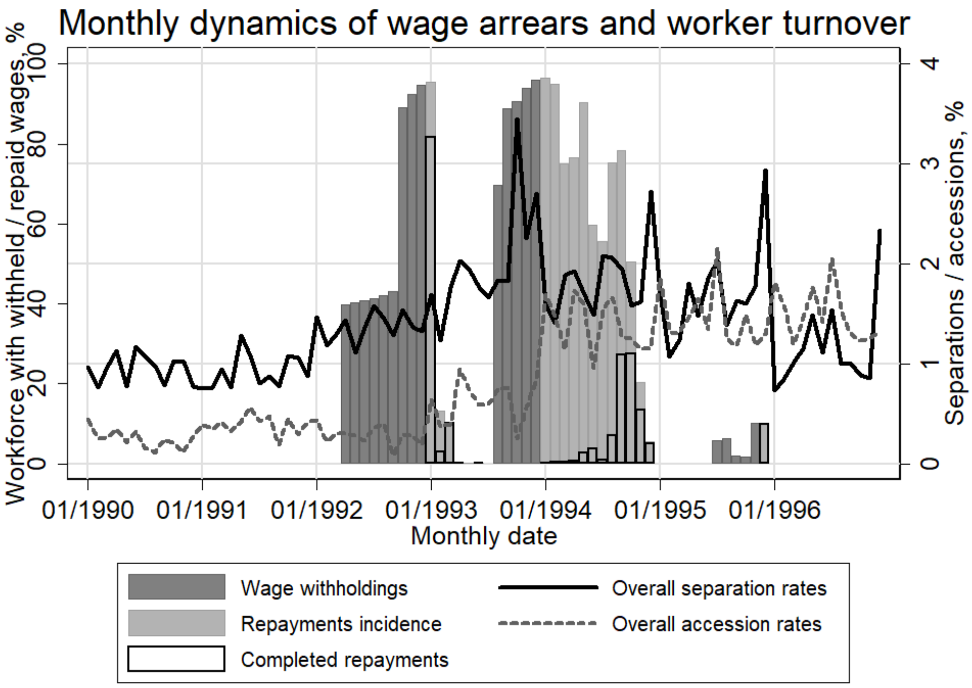


Figure IV.5: Separations, inflows, wage withholdings, and repayments

level, which is in line with the finding that the firm aims to offset employment changes triggered by workers' separation decisions.

IV.4 Concluding Discussion

We analyzed wage arrears under the aspect of deferred compensation, using unique personnel data from a large Russian manufacturing firm that straddle 1990 to 2006. We find evidence of strong bonding effects of deferred compensation during the two major episodes of wage arrears in this firm in 1992 to 1995. Having precise information on the timing and amount of nominal wage arrears and of repayments for each worker as well as on separation dates, we could test whether separation rates change during periods of outstanding wages and repayments. During the first episode of wage arrears our results show very strong bonding effects during periods for workers with outstanding wages and no change in turnover behavior during the repayment period.

Since these repayments were executed in nominal terms in times when monthly inflation oscillated between 10% and 40%, workers experienced large real wage losses when their compensation were deferred. We moot that the firm's reputation is suffering among workers as a consequence and that the deterioration of the firm's reputation should weaken the bonding effect when wage arrears occur again. We find that the bonding effect for workers affected by wage arrears in the second episode merely weakened. However, at the same time, the separation rate for the entire workforce increases, which indicates that the re-appearance of wage arrears pushes some workers to separate from the firm. This effect seems to be particularly high for workers who had been affected by wage arrears in the first period, but also workers not affected by wage arrears in the first round separate preventively from the firm, namely they leave the firm before they become potential victims of wage arrears themselves.

The following observation suggests that the firm was unable to pay wages and did not hide this fact from its workers. During the second episode of wage arrears, more than 25% of all workers were on unpaid leave: in September and October 1993, as well as in March–April and June–August 1994. The average length of these episodes of unpaid leave was 2 months in autumn 1993 and 2.65 months in 1994. The distribution of these episodes among different occupations was quite uniform: nearly one-third of the workforce from all six primary job occupations were on unpaid leave in autumn 1993 (much less in 1994). 9% of all workers (including managers) were on unpaid leave more than once.

The fact that wage arrears disappeared soon after they caused an increase in workers' turnover suggests that even if the firm might have taken advantage of wage arrears in earlier periods, it could not continue to employ this function in later ones.[29] This result excludes a popular view on wage arrears as a vicious cycle of "deviant practices", at least in case of our firm, which, of course, may only represent relatively successful manufacturing firms that were able to survive the transition, and not the average Russian firm. Our findings are most closely related to the idea of sustaining continuous employment for workers. As previously mentioned, more than one-quarter of the workforce has been on unpaid leave for 2 to 3 months during the second episode of wage arrears. These episodes of unpaid leave resulted in decreasing hazards shortly after they took place, holding the rest equal. This effect fades out over time, probably because the relative attractiveness of other options to workers has increased in later periods as the economy stabilizes.

Finally, the fact that workers tried to keep their workplace instead of trying to find employment elsewhere, changing occupation, or engaging in suit-case trading, can teach modern firms a valuable lesson. It is vitally important for a firm to comply with existing agreements on when it has to pay and what to pay to its workers to maintain long-term employment relationships. This fact also points to the greatest limitation of the study: since we do not observe workers' outside options, we cannot distinguish bonding effects from the absence of outside options.

[29] Less than 10% of workers were affected in 1995, and no wage arrears appear after 1995.

APPENDICES

IV.A Additional Results

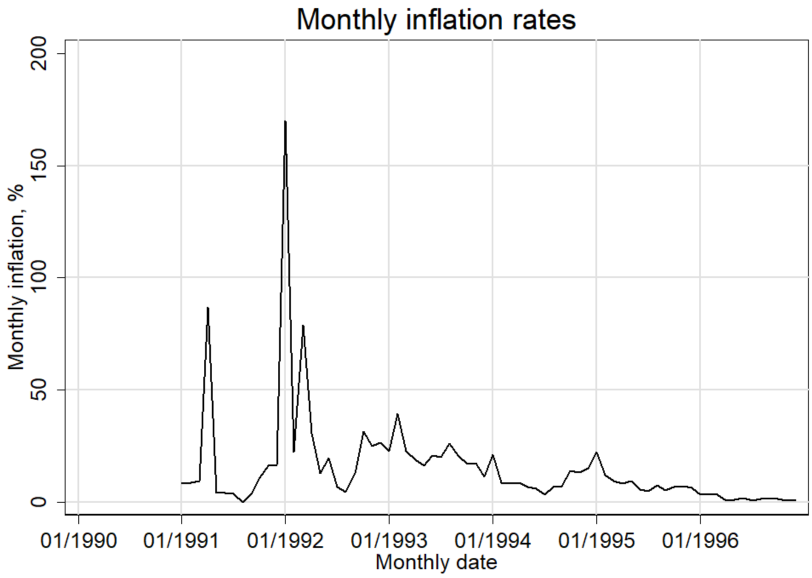


Figure IV.A.1: The dynamics of inflation

Source: Russian Statistical Agency.

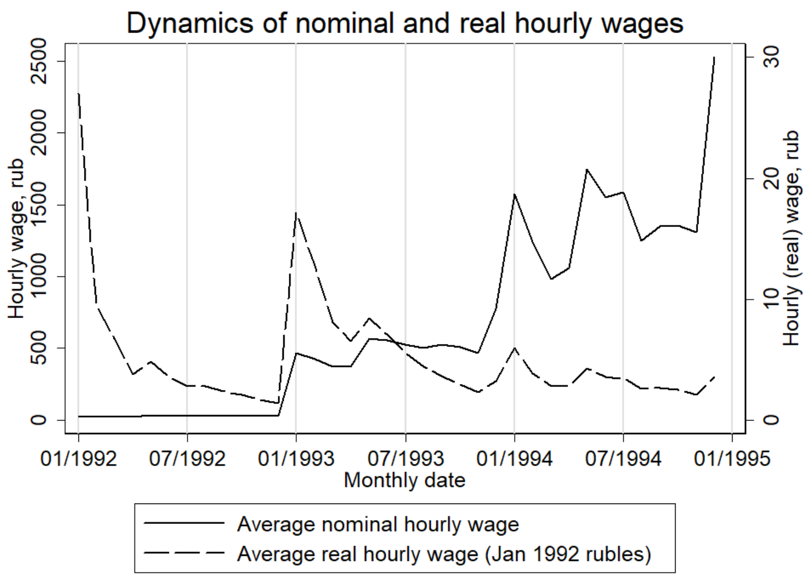


Figure IV.A.2: Dynamics of hourly wages in 1992–1994

Source: own computations.

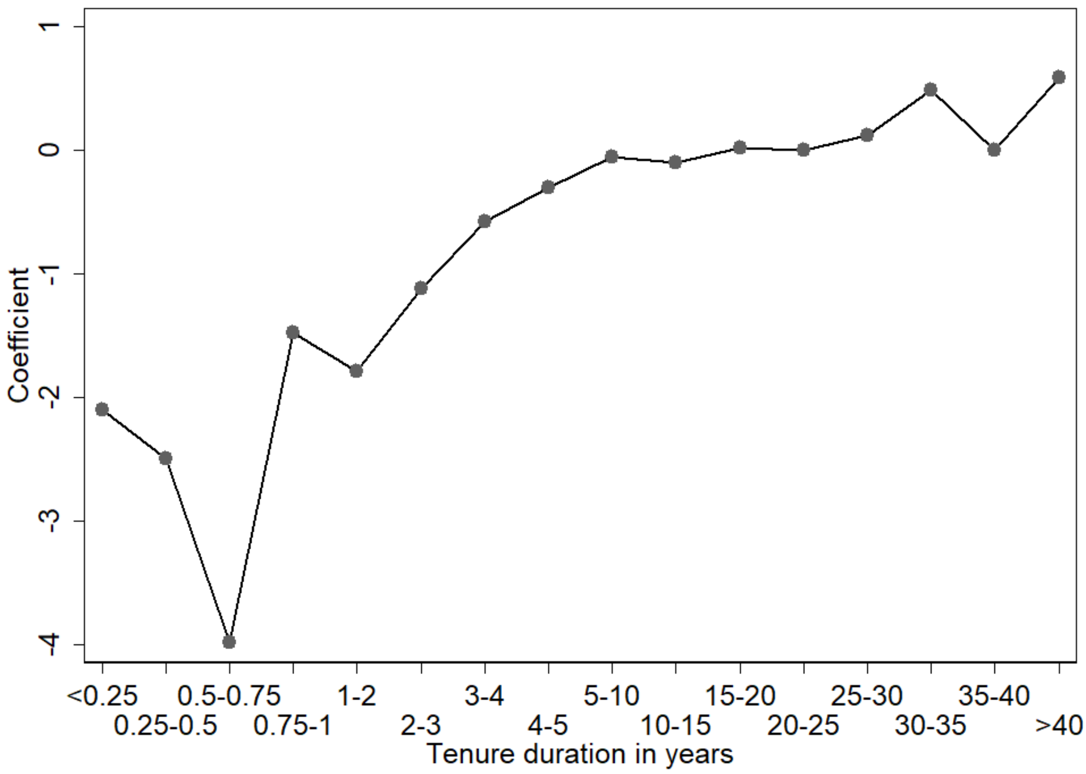


Figure IV.A.3: Coefficient estimates for tenure variables in a model, presented in Table 3, relative to the reference group (age=48)

Source: own computations.

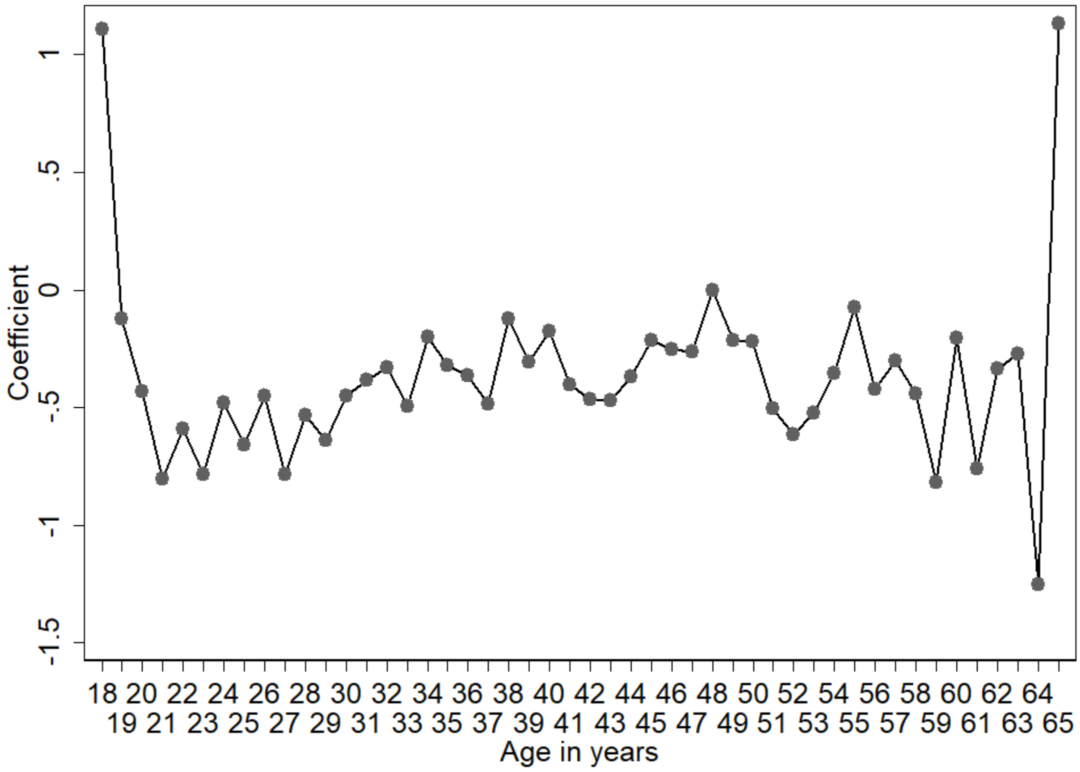


Figure IV.A.4: Coefficient estimates for age variables in a model presented in Table 3, relative to the reference group (age=48)

Source: own computations.

Table IV.A.1: Composition of the workforce

	Accountants	Managers	Engineers	Production workers		Service staff	Total # of	Real Wage
	˜	˜	˜	primary	auxiliary	˜	employees	(in Jan'92 rub.)
Year								
1991	1.1	4.9	29	43	20	8.3	4262	1573
1992	1.4	5.1	28	45	21	8.3	3881	681
1993	1.6	5.9	29	47	19	8.8	3393	819
1994	1.8	7	30	49	20	8.4	3029	337
1995	1.8	6.7	31	48	20	8.1	2896	313
1996	2	6.9	30	48	19	8.3	2900	193
1997	2.2	6.3	26	48	20	7.9	2964	466
1998	2	6.6	25	50	19	7.7	2991	532
1999	1.8	6.4	25	48	19	7	2955	301
2000	1.7	6.4	25	48	17	7	2975	300
2001	1.9	6.3	24	49	18	6.9	3042	292
2002	1.7	6.2	24	49	19	6.7	3064	305
2003	1.8	6.3	23	48	19	7.1	3084	352
2004	1.8	6.1	25	49	18	6.5	3204	350
2005	1.8	5.7	25	49	19	6.2	3403	381
2006	1.8	5.4	25	48	20	5.9	3618	382

Source: own computations.

Table IV.A.2: Number and percentage of workers whose wages were withheld, by month

Category	Statistic	Year	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Accountants	# of affected	1992	35	35	36	36	36	36	42	42	45
	percentage	1992	76.1	74.5	76.6	78.3	76.6	76.6	89.4	89.4	95.7
Managers	# of affected	1992	20	21	20	20	21	22	185	187	187
	percentage	1992	10.4	10.8	10.3	10.4	10.7	11.5	96.4	96.4	96.9
Engineers	# of affected	1992	831	844	832	833	834	833	895	898	905
	percentage	1992	79.1	80.2	80.3	81.1	82.3	83.1	90.7	92.1	93.9
Primary production workers	# of affected	1992	138	139	138	139	138	138	1436	1440	1450
	percentage	1992	8.38	8.42	8.60	8.71	8.76	8.86	93.1	94.4	96.2
Auxiliary production workers	# of affected	1992	543	547	546	546	546	546	576	579	591
	percentage	1992	74.2	75.8	76.9	78.4	80.1	82.0	87.7	89.6	93.2
Service staff	# of affected	1992	11	14	18	21	26	38	170	244	260
	percentage	1992	3.65	4.62	6.04	7.17	8.93	13.2	59.0	86.2	92.5
Accountants	# of affected	1993	.	.	.	.	39	44	44	44	44
	percentage	1993	.	.	.	.	81.3	95.7	97.8	97.8	97.8
Managers	# of affected	1993	.	.	.	.	160	186	185	190	186
	percentage	1993	.	.	.	.	82.5	93.9	95.4	96.0	97.4
Engineers	# of affected	1993	.	.	.	.	587	762	761	767	773
	percentage	1993	.	.	.	.	67.6	88.7	90.7	92.4	94.5
Primary production workers	# of affected	1993	.	.	.	.	985	1255	1263	1265	1279
	percentage	1993	.	.	.	.	70.9	91.3	92.2	95.0	97.0
Auxiliary production workers	# of affected	1993	.	.	.	.	387	469	475	477	481
	percentage	1993	.	.	.	.	68.0	84.1	87.0	92.1	95.4
Service staff	# of affected	1993	.	.	.	.	158	212	214	254	216
	percentage	1993	.	.	.	.	60.5	81.9	83.9	93.4	95.2
Accountants	# of affected	1995	.	.	.	3	2	2	.	3	.
	percentage	1995	.	.	.	5.88	3.92	3.92	.	5.88	.
Engineers	# of affected	1995	.	.	.	49	53	18	16	83	.
	percentage	1995	.	.	.	6.46	7.14	2.42	2.19	11.4	.
Primary production workers	# of affected	1995	.	.	.	76	80	25	26	129	.
	percentage	1995	.	.	.	6.16	6.51	2.03	2.11	10.5	.
Auxiliary production workers	# of affected	1995	.	.	.	22	31	9	2	46	.
	percentage	1995	.	.	.	4.55	6.44	1.88	0.42	9.70	.
Service staff	# of affected	1995	.	.	.	15	17	4	6	26	.
	percentage	1995	.	.	.	7.43	8.59	2.02	3.05	13.3	.

Source: own computations.

Table IV.A.3: Number and percentage of workers whose withheld wages were repaid, by month

Category	Statistic	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Accountants	# of fully repaid	1993	32	7	6	.	.	.	.	.	.	.	.	.
	percentage	1993	71.1	15.6	13.3	.	.	.	.	.	.	.	.	.
Managers	# of fully repaid	1993	175	6	.	.	.	.	.	.	.	.	.	.
	percentage	1993	93.6	3.17	.	.	.	.	.	.	.	.	.	.
Engineers	# of fully repaid	1993	811	29	43	.	.	3	.	.	.	.	.	.
	percentage	1993	89.6	3.25	4.89	.	.	0.36	.	.	.	.	.	.
Primary production workers	# of fully repaid	1993	1369	38	36	1	.	.	.	.	.	.	.	.
	percentage	1993	94.4	2.67	2.58	0.073	.	.	.	.	.	.	.	.
Auxiliary production workers	# of fully repaid	1993	540	17	30	1	.	.	.	.	.	.	.	.
	percentage	1993	91.4	2.93	5.22	0.18	.	.	.	.	.	.	.	.
Service staff	# of fully repaid	1993	15	5	240	.	.	.	.	.	.	.	.	.
	percentage	1993	5.77	1.95	94.1	.	.	.	.	.	.	.	.	.
Accountants	# of fully repaid	1994	.	.	.	1	1	4	.	1	19	7	6	3
	percentage	1994	.	.	.	2.33	2.33	9.30	.	2.63	50	18.4	16.2	8.82
Managers	# of fully repaid	1994	.	.	.	1	7	4	2	17	62	58	16	.
	percentage	1994	.	.	.	0.53	3.78	2.16	1.10	9.66	35.8	33.7	9.36	.
Engineers	# of fully repaid	1994	.	3	1	6	21	30	7	50	215	232	86	48
	percentage	1994	.	0.39	0.13	0.81	2.83	4.20	1	7.26	31.9	35.2	13.4	7.63
Primary production workers	# of fully repaid	1994	.	5	8	3	30	48	11	92	350	366	203	67
	percentage	1994	.	0.39	0.64	0.25	2.47	4.05	0.94	8.05	30.9	33.4	18.8	6.32
Auxiliary production workers	# of fully repaid	1994	1	2	1	8	20	20	5	43	120	108	59	23
	percentage	1994	0.21	0.38	0.22	1.80	4.48	4.82	1.23	10.9	31.6	29.4	16.4	6.59
Service staff	# of fully repaid	1994	.	.	2	1	8	6	1	10	47	46	30	7
	percentage	1994	.	.	0.93	0.48	3.94	3.05	0.51	5.29	25.4	25.4	17.0	4.02
Accountants	# of fully repaid	1995	.	.	.	.	.	.	.	.	.	.	.	3
	percentage	1995	.	.	.	.	.	.	.	.	.	.	.	8.82
Engineers	# of fully repaid	1995	.	.	.	.	.	.	.	.	.	.	.	83
	percentage	1995	.	.	.	.	.	.	.	.	.	.	.	18.5
Primary production workers	# of fully repaid	1995	.	.	.	.	.	.	.	.	.	.	.	126
	percentage	1995	.	.	.	.	.	.	.	.	.	.	.	14.3
Auxiliary production workers	# of fully repaid	1995	.	.	.	.	.	.	.	.	.	.	.	46
	percentage	1995	.	.	.	.	.	.	.	.	.	.	.	18.0
Service staff	# of fully repaid	1995	.	.	.	.	.	.	.	.	.	.	.	26
	percentage	1995	.	.	.	.	.	.	.	.	.	.	.	20.2

Source: own computations.

Table IV.A.4: Tenure variables

Dep. var.: $\lambda(t)$	Episode 1 Apr'92–Jul'93	Episode 2 Aug'93–Jan'95	Episode 2 Aug'93–Jan'95
Tenure 0 – 3 months	-2.307**	-4.908***	-2.302
	(0.747)	(0.458)	(1.446)
Tenure 3 – 6 months	-1.640**	-5.364***	-2.900
	(0.595)	(0.743)	(1.677)
Tenure 6 – 9 months	-45.297	-5.114***	-2.955*
	(.)	(1.058)	(1.494)
Tenure 9 – 12 months	0.010	-2.310***	-1.084*
	(0.497)	(0.527)	(0.524)
Tenure 1 – 2 years	-1.262***	-2.213***	-2.065***
	(0.376)	(0.364)	(0.366)
Tenure 2 – 3 years	-0.488	-2.009***	-2.005***
	(0.302)	(0.363)	(0.363)
Tenure 3 – 4 years	0.019	-1.361***	-1.358***
	(0.214)	(0.285)	(0.285)
Tenure 4 – 5 years	0.140	-0.557**	-0.553**
	(0.211)	(0.206)	(0.206)
Tenure 5 – 10 years	0.250	-0.263*	-0.260*
	(0.159)	(0.126)	(0.126)
Tenure 10 – 15 years	0.019	-0.250*	-0.249*
	(0.156)	(0.120)	(0.120)
Tenure 15 – 20 years	0.191	-0.128	-0.127
	(0.147)	(0.114)	(0.114)
Tenure 25 – 30 years	0.409*	-0.177	-0.178
	(0.192)	(0.182)	(0.182)
Tenure 30 – 35 years	0.245	0.223	0.222
	(0.352)	(0.241)	(0.242)
Tenure 35 – 40 years	0.201	-0.400	-0.399
	(0.604)	(0.477)	(0.477)
Tenure $\geq$ 40 years	0.955	0.215	0.215
	(0.731)	(0.643)	(0.644)
Other controls	Yes	Yes	Yes
Observations	58569	55253	55253

Notes. Cox proportional hazard estimates, robust standard errors in parentheses. Reported are tenure variables used in the model, presented in Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table IV.A.5: Age variables, pt. 1

Dep. var.: $\lambda(t)$	Episode 1 Apr'92–Jul'93	Episode 2 Aug'93–Jan'95	Episode 2 Aug'93–Jan'95
Age = 17	-40.620	-43.248	-43.066
	(.)	(.)	(.)
Age = 18	-43.542	0.670	0.808
	(.)	(1.214)	(1.210)
Age = 19	-45.348	0.212	0.410
	(.)	(1.104)	(1.077)
Age = 20	-0.847	-0.376	-0.349
	(0.647)	(0.827)	(0.819)
Age = 21	-1.624**	-0.180	-0.155
	(0.596)	(0.506)	(0.512)
Age = 22	-1.188**	-0.874*	-0.898*
	(0.412)	(0.432)	(0.434)
Age = 23	-1.029**	-0.987*	-1.002*
	(0.373)	(0.429)	(0.428)
Age = 24	-0.682*	-0.973*	-0.976*
	(0.316)	(0.399)	(0.398)
Age = 25	-1.479***	-0.846*	-0.855*
	(0.372)	(0.344)	(0.344)
Age = 26	-0.741*	-1.004**	-1.015**
	(0.305)	(0.321)	(0.321)
Age = 27	-0.939**	-1.048**	-1.060**
	(0.316)	(0.334)	(0.334)
Age = 28	-0.866**	-0.608*	-0.619*
	(0.296)	(0.289)	(0.288)
Age = 29	-0.944**	-1.016***	-1.026***
	(0.300)	(0.299)	(0.299)
Age = 30	-0.554	-0.603*	-0.603*
	(0.284)	(0.272)	(0.273)
Age = 31	-0.749**	-0.375	-0.379
	(0.290)	(0.283)	(0.283)
Age = 32	-0.635*	-0.362	-0.364
	(0.275)	(0.240)	(0.240)
Age = 33	-0.431	-0.737**	-0.736**
	(0.272)	(0.257)	(0.257)
Age = 34	-0.217	-0.295	-0.296
	(0.263)	(0.233)	(0.233)
Age = 35	-0.507	-0.210	-0.211
	(0.275)	(0.243)	(0.243)
Age = 36	-0.364	-0.271	-0.272
	(0.269)	(0.246)	(0.245)
Age = 37	-0.593*	-0.512*	-0.512*
	(0.286)	(0.241)	(0.241)
Age = 38	-0.054	-0.299	-0.299
	(0.259)	(0.233)	(0.233)
Age = 39	-0.249	-0.277	-0.280
	(0.275)	(0.236)	(0.236)
Age = 40	-0.206	-0.200	-0.200
	(0.261)	(0.234)	(0.234)
Other controls	Yes	Yes	Yes
Observations	58569	55253	55253

Notes. Cox proportional hazard estimates, robust standard errors in parentheses. Reported are age variables used in the model, presented in Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table IV.A.6: Age variables, pt. 2

Dep. var.: $\lambda(t)$	Episode 1 Apr'92–Jul'93	Episode 2 Aug'93–Jan'95	Episode 2 Aug'93–Jan'95
Age = 41	-0.380	-0.679*	-0.679*
	(0.280)	(0.269)	(0.269)
Age = 42	-0.565	-0.515*	-0.519*
	(0.294)	(0.263)	(0.263)
Age = 43	-0.130	-0.514*	-0.514*
	(0.274)	(0.238)	(0.238)
Age = 44	-0.331	-0.284	-0.283
	(0.283)	(0.234)	(0.235)
Age = 45	-0.065	-0.133	-0.135
	(0.273)	(0.237)	(0.238)
Age = 46	-0.225	-0.293	-0.293
	(0.304)	(0.251)	(0.252)
Age = 47	-0.059	-0.268	-0.268
	(0.317)	(0.247)	(0.248)
Age = 49	-0.148	-0.344	-0.344
	(0.331)	(0.315)	(0.315)
Age = 50	-0.312	0.215	0.217
	(0.358)	(0.269)	(0.269)
Age = 51	-0.474	-0.252	-0.250
	(0.348)	(0.271)	(0.272)
Age = 52	-0.632	-0.446	-0.448
	(0.371)	(0.295)	(0.296)
Age = 53	-0.271	-0.455	-0.457
	(0.435)	(0.302)	(0.303)
Age = 54	-0.034	-0.450	-0.452
	(0.367)	(0.377)	(0.379)
Age = 55	-0.213	0.321	0.320
	(0.434)	(0.317)	(0.317)
Age = 56	-0.794	-0.336	-0.338
	(0.492)	(0.380)	(0.380)
Age = 57	-0.501	-0.571	-0.573
	(0.545)	(0.335)	(0.335)
Age = 58	-0.059	-0.453	-0.450
	(0.446)	(0.561)	(0.561)
Age = 59	-0.250	-0.590	-0.588
	(0.606)	(0.516)	(0.516)
Age = 60	-0.715	-0.391	-0.396
	(0.754)	(0.609)	(0.610)
Age = 61	-1.654	-0.152	-0.155
	(1.022)	(0.605)	(0.606)
Age = 62	-0.604	-0.801	-0.804
	(0.729)	(0.541)	(0.541)
Age = 63	-0.268	-0.959	-0.964
	(0.986)	(0.710)	(0.709)
Age = 64	-45.117	-0.302	-0.310
	(.)	(1.223)	(1.222)
Age = 65	0.270	1.344	1.340
	(0.738)	(0.824)	(0.827)
Other controls	Yes	Yes	Yes
Observations	58569	55253	55253

Notes. Cox proportional hazard estimates, robust standard errors in parentheses. Reported are age variables used in the model, presented in Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table IV.A.7: Position variables, pt. 1

Dep. var.: $\lambda(t)$	Episode 1 Apr'92–Jul'93	Episode 2 Aug'93–Jan'95	Episode 2 Aug'93–Jan'95
1 if got promotion	0.083	-1.379***	-1.380***
	(0.426)	(0.398)	(0.398)
1 if got demotion	-41.459	-0.980	-0.978
	(.)	(0.600)	(0.600)
1 if got horizontal change	0.004	-0.479**	-0.482**
	(0.485)	(0.160)	(0.161)
Accountant	0.693	0.321	0.315
	(0.428)	(0.274)	(0.274)
Chief accountant	-44.439	-45.120	-45.146
	(.)	(.)	(.)
Manager	-1.146***	-1.064***	-1.060***
	(0.294)	(0.231)	(0.231)
Technician/Engineer	0.490**	-0.562***	-0.558***
	(0.181)	(0.136)	(0.137)
Technician/Engineer, cat. 1	0.796**	-0.654*	-0.664*
	(0.259)	(0.302)	(0.302)
Technician/Engineer, cat. 2	0.214	-0.837***	-0.835***
	(0.288)	(0.251)	(0.251)
Technician/Engineer, cat. 3	0.760	-0.133	-0.156
	(0.501)	(1.150)	(1.132)
Chief technician/Chief engineer	0.234	-0.598*	-0.594*
	(0.312)	(0.262)	(0.262)
Other controls	Yes	Yes	Yes
Observations	58569	55253	55253

Notes. Cox proportional hazard estimates, robust standard errors in parentheses. Reported are position variables used in the model, presented in Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table IV.A.8: Position variables, pt. 2

Dep. var.: $\lambda(t)$	Episode 1 Apr'92–Jul'93	Episode 2 Aug'93–Jan'95	Episode 2 Aug'93–Jan'95
Primary production worker	-44.712	-2.410*	-2.409*
	(.)	(0.941)	(0.941)
Primary production worker, cat. 1	1.123*	-0.162	-0.118
	(0.546)	(1.071)	(1.080)
Primary production worker, cat. 2	0.822***	-0.017	-0.004
	(0.219)	(0.206)	(0.208)
Primary production worker, cat. 3	0.364*	-0.137	-0.137
	(0.151)	(0.119)	(0.119)
Primary production worker, cat. 4	0.117	-0.219	-0.216
	(0.171)	(0.134)	(0.134)
Primary production worker, cat. 5	-0.654**	-0.873***	-0.872***
	(0.220)	(0.159)	(0.160)
Primary production worker, cat. 6	-0.930**	-1.108***	-1.106***
	(0.340)	(0.241)	(0.241)
Primary production worker, cat. 7	-0.423	-0.116	-0.107
	(0.946)	(0.577)	(0.577)
Primary production worker, cat. 8	-42.263	-43.981	-44.011
	(.)	(.)	(.)
Auxiliary production worker	-42.072	-43.241	-43.258
	(.)	(.)	(.)
Auxiliary production worker, cat. 1	2.034	1.629*	1.841**
	(1.112)	(0.815)	(0.707)
Auxiliary production worker, cat. 2	0.970***	0.592**	0.604**
	(0.271)	(0.227)	(0.226)
Auxiliary production worker, cat. 3	1.205***	0.254	0.257
	(0.167)	(0.138)	(0.138)
Auxiliary production worker, cat. 4	0.824***	-0.030	-0.031
	(0.199)	(0.169)	(0.170)
Auxiliary production worker, cat. 5	0.338	-0.323	-0.320
	(0.276)	(0.191)	(0.191)
Auxiliary production worker, cat. 6	-44.728	-1.107*	-1.103*
	(.)	(0.510)	(0.510)
Service worker	0.000	0.000	0.000
	(.)	(.)	(.)
Other controls	Yes	Yes	Yes
Observations	58569	55253	55253

Notes. Cox proportional hazard estimates, robust standard errors in parentheses. Reported are position variables used in the model, presented in Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table IV.A.9: Basic individual controls

Dep. var.: $\lambda(t)$	Episode 1 Apr'92–Jul'93	Episode 2 Aug'93–Jan'95	Episode 2 Aug'93–Jan'95
1 if has higher education	0.320*	0.316*	0.314*
	(0.137)	(0.136)	(0.136)
1 if incomplete higher	-0.105	0.143	0.110
	(0.415)	(0.439)	(0.440)
1 if complete professional	0.144	0.376***	0.376***
	(0.079)	(0.077)	(0.077)
1 if female	-0.182*	-0.163*	-0.162*
	(0.087)	(0.081)	(0.082)
1 if married	-1.497***	-1.493***	-1.484***
	(0.337)	(0.347)	(0.348)
1 if has kids	-0.687***	-0.400***	-0.403***
	(0.083)	(0.088)	(0.088)
1 if of retire age	0.816*	0.216	0.219
	(0.402)	(0.364)	(0.364)
Other controls	Yes	Yes	Yes
Observations	58569	55253	55253

Notes. Cox proportional hazard estimates, robust standard errors in parentheses. Reported are individual characteristics variables used in the model, presented in Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

IV.B Wage Growth and the Likelihood of Position Changes

IV.B.1 Likelihood of Position Changes in 1992–1995

Table IV.B.1: Fixed Effects Logit for Mobility

	Dependent variable		
	Promotion (1)	Demotion (2)	Horizontal change (3)
1 if wage withheld in $t-1$	0.730	17.451**	2.721***
	(0.199)	(15.330)	(0.472)
1 if repayment happened in $t-1$	2.095***	5.601*	3.670***
	(0.392)	(4.120)	(0.515)
Hourly wage in $t-1$	1.000	1.000	1.000
	(0.000)	(0.000)	(0.000)
Age (years)	1.495	0.186	1.124
	(0.612)	(0.243)	(0.341)
Age^2 ($\text{years}^2/100$)	0.626	10.877	0.947
	(0.271)	(16.421)	(0.299)
General tenure (years)	0.920	2.488	1.132
	(0.261)	(2.208)	(0.235)
1 if fully repaid in t	0.603	0.000	0.651
	(0.280)	(0.000)	(0.171)
Observations	6456	870	11552
Number of individuals	165	22	300
Model	fe	fe	fe

Notes. Fixed effects Logit estimates, reported are odds ratios. SeEform in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

IV.B.2 Likelihood of Being Promoted in One Year after January 1994

Table IV.B.2: Number of promoted workers among affected by wage arrears in Jan'94

	Affected vs Promoted Not affected	Affected	Total
Not promoted	46	2246	2292
Promoted	2	57	59
Total	48	2303	2351
Fisher's Exact test p	0.341		

Notes. Fisher's Exact test 2-sided p-value is reported.

IV.B.3 One-Year Wage Growth since January 1994

Table IV.B.3: One-year log-wage growth and the corresponding change in percentile ranking in wage distribution since Jan'94

Dep. var.: Jan'94–Jan'95 change in	(1) 1-year log-wage growth Log(Hourly wages)	(2) 1-year wage rank change Pct. in wage distr.
1 if affected in Jan'94	0.995***	-14.860
	(0.103)	(14.729)
1 if has higher education	0.096	0.977
	(0.063)	(3.421)
1 if incomplete higher	0.156	8.479
	(0.094)	(7.352)
1 if complete professional	0.056	2.678
	(0.040)	(2.139)
1 if female	0.034	0.341
	(0.037)	(2.168)
1 if married	-0.206*	-12.658**
	(0.095)	(4.750)
1 if has kids	0.005	-2.547
	(0.053)	(2.788)
1 if of retire age	0.359	11.605
	(0.196)	(12.017)
Constant	-0.427	2.774
	(0.219)	(19.314)
Other controls	Yes	Yes
Observations	2309	2324

Notes. Reported are OLS estimates. Other controls include age and tenure dummies, position dummies and dummies for position changes. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table IV.B.4: Position variables for specifications in Table IV.B.3

Dep. var.: Jan'94–Jan'95 change in	(1) 1-year log-wage growth Log(Hourly wages)	(2) 1-year wage rank change Pct. in wage distr.
1 if got promotion	-0.078	-18.807**
	(0.131)	(5.792)
1 if got demotion	0.084	20.919
	(0.125)	(14.118)
1 if got horizontal change	0.039	5.271
	(0.058)	(3.468)
Accountant	0.000	0.000
	(.)	(.)
Chief accountant	0.153	0.077
	(0.412)	(15.259)
Manager	0.393**	-3.029
	(0.133)	(10.146)
Technician/Engineer	0.144	-0.387
	(0.120)	(9.592)
Technician/Engineer, cat. 1	-0.023	-1.057
	(0.159)	(11.013)
Technician/Engineer, cat. 2	0.183	-0.041
	(0.142)	(10.520)
Technician/Engineer, cat. 3	0.200	-3.747
	(0.141)	(25.504)
Chief technician/Chief engineer	0.032	0.547
	(0.163)	(10.904)
Primary production worker	0.304*	22.900*
	(0.135)	(10.336)
Primary production worker, cat. 1	0.174	26.341
	(0.148)	(29.801)
Primary production worker, cat. 2	-0.086	-3.628
	(0.227)	(11.849)
Primary production worker, cat. 3	0.270*	0.489
	(0.130)	(9.944)
Primary production worker, cat. 4	0.336*	2.220
	(0.136)	(10.135)
Primary production worker, cat. 5	0.380**	-1.055
	(0.132)	(10.142)
Primary production worker, cat. 6	0.138	0.852
	(0.150)	(10.853)
Primary production worker, cat. 7	0.660*	4.097
	(0.309)	(26.587)
Primary production worker, cat. 8	0.209	-4.760
	(0.148)	(10.712)
Auxiliary production worker	0.123	-3.565
	(0.139)	(10.571)
Auxiliary production worker, cat. 2	0.069	8.723
	(0.137)	(12.439)
Auxiliary production worker, cat. 3	0.172	1.481
	(0.137)	(10.195)
Auxiliary production worker, cat. 4	0.078	-1.660
	(0.136)	(10.671)
Auxiliary production worker, cat. 5	0.209	0.599
	(0.138)	(10.622)
Auxiliary production worker, cat. 6	0.222	-5.345
	(0.232)	(16.888)
Service worker	0.507***	3.266
	(0.135)	(10.221)
Other controls	Yes	Yes
Observations	2309	2324

Notes. Reported are OLS estimates. Other controls include age and tenure dummies, and variables reported in Table IV.B.3. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Bibliography

Kommersant, 2006. No. 182 (3513), 29.09.2006.

J. H. Abbring and G. J. Van den Berg. The Nonparametric Identification of Treatment Effects in Duration Models. *Econometrica*, 71(5):1491–1517, 2003.

G. A. Akerlof and L. F. Katz. Workers' Trust Funds and the Logic of Wage Profiles. *The Quarterly Journal of Economics*, 104(3):525–536, 1989.

M. Ananyev. Leading-by-Example in a Public Goods Game with Heterogeneous Returns. Technical report, 2019.

L. Anderson, J. Mellor, and J. Milyo. Inequality, Group Cohesion, and Public Good Provision: An Experimental Analysis. Technical report, 2004.

J. Andreoni and L. K. Gee. Gun for hire: Delegated enforcement and peer punishment in public goods provision. *Journal of Public Economics*, 96(11-12):1036–1046, 2012.

E. Arbak and M.-C. Villeval. Voluntary leadership: motivation and influence. *Social Choice and Welfare*, 40(3):635–662, 2013.

W. T. Au and M. W. Chung. Effect of order of contribution in a sequential public goods dilemma. *Group Decision and Negotiation*, 16(5):437–449, 2007.

G. S. Becker and G. J. Stigler. Law Enforcement, Malfeasance, and Compensation of Enforcers. *The Journal of Legal Studies*, 3(1):1–18, 1974.

O. Bock, I. Baetge, and A. Nicklisch. hroot: Hamburg registration and organization online tool. *European Economic Review*, 71:117–120, 2014.

G. E. Bolton and A. Ockenfels. Erc: A theory of equity, reciprocity, and competition. *The American Economic Review*, pages 166–193, 2000.

K. A. Brekke, K. E. Hauge, J. T. Lind, and K. Nyborg. Playing with the good guys. a public good game with endogenous group formation. *Journal of Public Economics*, 95 (9):1111–1118, 2011.

E. Buckley and R. Croson. Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics*, 90(4):935–955, 2006.

C. Bull. The Existence of Self-Enforcing Implicit Contracts. *The Quarterly Journal of Economics*, 102(1):147–159, 1987.

A. W. Cappelen, A. D. Hole, E. Ø. Sørensen, and B. Tungodden. The pluralism of fairness ideals: An experimental approach. *The American Economic Review*, 97(3): 818–827, 2007.

H. L. Carmichael. Self-Enforcing Contracts, Shirking, and Life Cycle Incentives. *Journal of Economic Perspectives*, 3(4):65–83, 1989.

S. Centorrino, L. Concina, et al. A competitive approach to leadership in public good games. Technical Report TSE-383, Toulouse School of Economics (TSE), 2013.

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

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

K. S. Chan, S. Mestelman, R. A. Muller, and R. Moir. Communication, equity and the voluntary provision of a public good by heterogeneous groups. *The Journal of Socio-Economics*, 41(1):87–94, 2012.

G. Charness and M. Rabin. Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3):817–869, 2002.

T. L. Cherry, S. Kroll, and J. F. Shogren. The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. *Journal of Economic Behavior & Organization*, 57(3):357–365, 2005.

D. R. Cox. Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202, 1972.

D. R. Cox. Partial Likelihood. *Biometrika*, 62(2):269–276, 1975.

J. C. Cox, E. Ostrom, and J. M. Walker. Bosses and kings: asymmetric power in paired common pool and public good games. In *Proceedings of the Behavioral and Quantitative Game Theory: Conference on Future Directions*, page 82. ACM, 2010.

J. Dana, R. A. Weber, and J. X. Kuang. Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1):67–80, 2007.

P. Desai and T. Idson. Wage arrears, poverty, and family survival strategies. *Work Without Wages: Russia's Nonpayment Crisis*, pages 207–223, 2000.

T. J. Dohmen and G. A. Pfann. Worker Separations in a Nonstationary Corporate Environment. *European Economic Review*, 48(3):645–663, 2004.

M. Drouvelis and D. Nosenzo. Group identity and leading-by-example. *Journal of Economic Psychology*, 39:414–425, 2013.

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

J. S. Earle and K. Z. Sabirianova. How Late to Pay? Understanding Wage Arrears in Russia. *Journal of Labor Economics*, 20(3):661–707, 2002.

J. S. Earle and K. Sabirianova Peter. Complementarity and Custom in Wage Contract Violation. *The Review of Economics and Statistics*, 91(4):832–849, 2009.

J. S. Earle, A. Spicer, and K. S. Peter. The Normalization of Deviant Organizational Practices: Wage Arrears in Russia, 1991–98. *Academy of Management Journal*, 53(2): 218–237, 2010.

K.-M. Ehrhart and C. Keser. Mobility and cooperation: On the run. Technical report, 1999.

I. Erev and A. Rapoport. Provision of step-level public goods: The sequential contribution mechanism. *Journal of Conflict Resolution*, 34(3):401–425, 1990.

A. Falk and U. Fischbacher. A theory of reciprocity. *Games and economic behavior*, 54 (2):293–315, 2006.

A. Falk and M. Knell. Choosing the joneses: endogenous goals and reference standards. *Scandinavian Journal of Economics*, 106(3):417–435, 2004.

E. Fehr and S. Gächter. Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4):980–994, 2000.

E. Fehr and S. Gächter. Altruistic punishment in humans. *Nature*, 415(6868):137–140, 2002.

E. Fehr and K. M. Schmidt. A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3):817–868, 1999.

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

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

U. Fischbacher, S. Schudy, and S. Teyssier. Heterogeneous reactions to heterogeneity in returns from public goods. *Social Choice and Welfare*, 43(1):195–217, 2014.

J. Fisher, R. M. Isaac, J. M. Walker, J. W. Schatzberg, and J. M. Walker. Heterogenous Demand for Public Goods : Behavior in the Voluntary Contributions Mechanism. *Public Choice*, 85(3):249–266, 1995.

G. Frackenpohl, A. Hillenbrand, and S. Kube. Leadership effectiveness and institutional frames. *Experimental Economics*, 19(4):842–863, 2016.

S. Gächter and E. Renner. Leaders as role models and 'belief managers' in social dilemmas. *Journal of Economic Behavior & Organization*, 154:321–334, 2018.

S. Gächter and C. Thöni. Social learning and voluntary cooperation among like-minded people. *Journal of the European Economic Association*, 3(2-3):303–314, 2005.

S. Gächter, D. Nosenzo, E. Renner, and M. Sefton. Sequential vs. simultaneous contributions to public goods: Experimental evidence. *Journal of Public Economics*, 94(7-8): 515–522, 2010.

S. Gächter, D. Nosenzo, E. Renner, and M. Sefton. Who makes a good leader? cooperativeness, optimism, and leading-by-example. *Economic Inquiry*, 50(4):953–967, 2012.

V. Gimpelson and R. Kapeliushnikov. Labor Market Adjustment: Is Russia Different? In *The Oxford Handbook of the Russian Economy*. 2013.

A. Glöckner, B. Irlenbusch, S. Kube, A. Nicklisch, and H.-T. Normann. Leading with (out) sacrifice? a public-goods experiment with a privileged player. *Economic Inquiry*, 49(2):591–597, 2011.

J. Gokhale, E. L. Groshen, and D. Neumark. Do Hostile Takeovers Reduce Extramarginal Wage Payments? *The Review of Economics and Statistics*, 77(3):470–485, 1995.

Ö. Gürerk, B. Irlenbusch, and B. Rockenbach. The competitive advantage of sanctioning institutions. *Science*, 312(5770):108–111, 2006.

Ö. Gürerk, B. Irlenbusch, and B. Rockenbach. Motivating teammates: The leader's choice between positive and negative incentives. *Journal of Economic Psychology*, 30(4):591–607, 2009.

W. Güth, M. V. Levati, M. Sutter, and E. Van Der Heijden. Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics*, 91(5):1023–1042, 2007.

S. D. Haigner and F. Wakolbinger. To lead or not to lead: Endogenous sequencing in public goods games. *Economics Letters*, 108(1):93–95, 2010.

J. R. Hamman, R. A. Weber, and J. Woon. An experimental investigation of electoral delegation and the provision of public goods. *American Journal of Political Science*, 55 (4):738–752, 2011.

B. Hartig, B. Irlenbusch, and F. Kölle. Conditioning on What? Heterogeneous Contributions and Conditional Cooperation. *Journal of Behavioral and Experimental Economics*, 55:48–64, 2015.

R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, and V. Gupta. *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Sage publications, 2004.

S. Huck, A. J. Seltzer, and B. Wallace. Deferred Compensation in Multiperiod Labor Contracts: An Experimental Test of Lazear's Model. *American Economic Review*, 101 (2):819–43, 2011.

M. Ibanez and E. Schaffland. Organizational performance with in-group and out-group leaders: An experiment. *Journal of behavioral and experimental economics*, 73:1–10, 2018.

G. W. Imbens. Transition Models in a Non-Stationary Environment. *The Review of Economics and Statistics*, 76(4):703–720, 1994.

M. C. Jensen. Leadership. *Harvard NOM Working Paper No. 06-05; Barbados Working Paper No. 05-04*, June 2005. Available at SSRN: https://ssrn.com/abstract=744786 or http://dx.doi.org/10.2139/ssrn.744786.

B. Jovanovic. Firm-Specific Capital and Turnover. *Journal of Political Economy*, 87(6): 1246–1260, 1979.

L. Kießling, J. Radbruch, and S. Schaube. The impact of self-selection on performance. Technical report, IZA Discussion Paper No 11365, 2018.

S. Kroll, T. L. Cherry, and J. F. Shogren. The impact of endowment heterogeneity and origin on contributions in best-shot public good games. *Experimental Economics*, 10 (4):411–428, 2007.

S. Kube, S. Schaube, H. Schildberg-Hörisch, and E. Khachatryan. Institution formation and cooperation with heterogeneous agents. *European Economic Review*, 78:248–268, 2015.

R. Layard and A. Richter. Labour Market Adjustment — the Russian Way. In A. Aslund, editor, *Russian Economic Reform at Risk*, pages 119–148. Pinter, London, UK, 1995a. ISBN 1855672863.

R. Layard and A. Richter. How Much Unemployment Is Needed for Restructing: The Russian Experience. *Economics of Transition*, 3(1):39–58, 1995b.

E. P. Lazear. Why Is There Mandatory Retirement? *Journal of Political Economy*, 87 (6):1261–1284, 1979.

J. Ledyard. Public goods: A survey of experimental research. In J. H. Kagel and A. E. Roth, editors, *The handbook of experimental economics*. Princeton university press Princeton, NJ, Princeton, 1995.

H. Lehmann and J. Wadsworth. Tenures That Shook the World: Worker Turnover in Russia, Poland, and Britain. *Journal of Comparative Economics*, 28(4):639–664, 2000.

H. Lehmann, J. Wadsworth, and A. Acquisti. Grime and Punishment: Job Insecurity and Wage Arrears in the Russian Federation. *Journal of Comparative Economics*, 27(4): 595–617, 1999.

M. V. Levati, M. Sutter, and E. Van der Heijden. Leading by example in a public goods experiment with heterogeneity and incomplete information. *Journal of Conflict Resolution*, 51(5):793–818, 2007.

D. M. Levy, K. Padgitt, S. J. Peart, D. Houser, and E. Xiao. Leadership, cheap talk and really cheap talk. *Journal of Economic Behavior & Organization*, 77(1):40–52, 2011.

W. B. MacLeod and J. M. Malcomson. Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment. *Econometrica: Journal of the Econometric Society*, 57(2): 447–480, 1989.

W. B. MacLeod and J. M. Malcomson. Specific Investment and Wage Profiles in Labour Markets. *European Economic Review*, 37(2-3):343–354, 1993.

W. B. MacLeod and J. M. Malcomson. Motivation and Markets. *American Economic Review*, 88(3):388–411, 1998.

B. Monin and D. T. Miller. Moral credentials and the expression of prejudice. *Journal of personality and social psychology*, 81(1):33, 2001.

J. Neitzel and L. Sääksvuori. Normative Conflict and Cooperation in Sequential Social Dilemmas. Technical report, 2013.

N. Nikiforakis, C. N. Noussair, and T. Wilkening. Normative conflict and feuds: The limits of self-enforcement. *Journal of Public Economics*, 96(9):797–807, 2012.

N. Nikiforakis, M. C. Villeval, and L. Gangadharan. Equality Concerns and the Limits of Self-Governance in Heterogeneous Populations. Technical Report 9384, 2015.

C. N. Noussair and F. Tan. Voting on punishment systems within a heterogeneous group. *Journal of Public Economic Theory*, 13(5):661–693, 2011.

R. J. Oxoby and J. Spraggon. A Clear and Present Minority : Heterogeneity in the Source of Endowments and the Provision of Public Goods. Technical report, 2006.

J. Potters, M. Sefton, L. Vesterlund, et al. Why announce leadership contributions?: An experimental study of the signaling and reciprocity hypotheses. Technical report, 2001.

J. Potters, M. Sefton, and L. Vesterlund. After you—endogenous sequencing in voluntary contribution games. *Journal of public Economics*, 89(8):1399–1419, 2005.

J. Potters, M. Sefton, and L. Vesterlund. Leading-by-example and signaling in voluntary contribution games: an experimental study. *Economic Theory*, pages 169–182, 2007.

E. Reuben and A. Riedl. Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior*, 77(1):122–137, 2013.

M. F. Rivas and M. Sutter. The benefits of voluntary leadership in experimental public goods games. *Economics Letters*, 112(2):176–178, 2011.

T. C. Schelling. *The strategy of conflict*. Harvard University Press, Cambridge, MA, 1960.

R. Selten. Die strategiemethode zur erforschung des engeschränkt rationalen verhaltens im rahmen eines oligopolexperiments. In H. Sauermann, editor, *Beiträge zur Experimentellen Wirtschaftsforschung*, pages 136–168. J. C. B. Mohr, Tübingen, 1967.

F. Tan. Punishment in a linear public good game with productivity heterogeneity. *De Economist*, 156(3):269–293, 2008.

E. Van der Heijden, J. Potters, and M. Sefton. Hierarchy and opportunism in teams. *Journal of Economic Behavior & Organization*, 69(1):39–50, 2009.

J. J. van der Weele, J. Kulisa, M. Kosfeld, and G. Friebel. Resisting moral wiggle room: how robust is reciprocal behavior? *American economic Journal: microeconomics*, 6(3): 256–64, 2014.

E. Van Dijk and H. Wilke. Differential interests, equity, and public good provision. *Journal of Experimental Social Psychology*, 29(1):1–16, 1993.

E. Van Dijk and H. Wilke. Coordination rules in asymmetric social dilemmas: A comparison between public good dilemmas and resource dilemmas. *Journal of Experimental Social Psychology*, 31(1):1–27, 1995.

E. Van Dijk and H. Wilke. Decision-induced focusing in social dilemmas: give-some, keep-some, take-some, and leave-some dilemmas. *Journal of personality and social psychology*, 78(1):92, 2000.

H.-M. von Gaudecker. Templates for reproducible research projects in economics. https://github.com/hmgaudecker/econ-project-templates, 2014.

D. A. Wise. *Pensions, Labor, and Individual Choice*. University of Chicago Press, 1985.

T. Yamakawa, Y. Okano, and T. Saijo. Detecting motives for cooperation in public goods experiments. *Experimental Economics*, 19(2):500–512, 2016.

H. P. Young. Social norms. In S. N. Durlauf, L. Blume, et al., editors, *The new Palgrave dictionary of economics*, volume 7, pages 647–651. Palgrave Macmillan, Basingstoke, 2008.