Essays on Behavioral and Experimental Economics
Cooperation, Emotions, and Health

Emil Persson

UNIVERSITY OF
GOTHENBURG

To Anna, Sara and Tilda

Acknowledgments

First and foremost I want to thank my supervisor Peter Martinsson. I have benefited tremendously from his support and insightful advice. Peter has been an inspiration and an academic role model. Martin Kocher and Gustav Tinghög also played a huge role in the development of this thesis. They kindly invited me to their labs and provided valuable comments and suggestions for some or all of my papers.

I would also like to thank Martin Dufwenberg and Amrish Patel for inspiring discussions and very helpful advice on Paper III, and Olof Johansson-Stenman, Gustav Kjellson, Dirk Engelmann, Abigail Barr, and Conny Wollbrant for valuable comments on my work at important stages during my PhD journey. Special thanks to Andrea Martinangeli and Lisa Björk who provided invaluable assistance for the experiment in Paper I. I am also grateful to Joakim Gullstrand, who encouraged me to apply to the PhD program.

Thanks to everyone at the Department of Economics at the University of Gothenburg. Many of you have helped me and inspired me during my time here.

I would like to give my deepest thanks to my family. To my mum Harriet and my brother Gustav: you have always supported me and shown great interest in what I do. Above all, I am grateful to my wife Anna and our two daughters Sara and Tilda for your love and all the joy. Anna, you have been a source of constant support and encouragement – this journey was difficult at times and you helped me through it.

Emil Persson
Gothenburg, January 2016

Contents

iv

1 Introduction

This thesis explores issues in human behavior and their implications for policy design. In four self-contained papers, I introduce new aspects in the analysis of cooperation and health provision, and I also investigate the strategic consequences of emotions (frustration and anger) in human interaction. These topics are investigated using laboratory experiments, a method that facilitates both variation in key institutional features to allow causal inference and analysis of sharp theoretical predictions in a clean environment.

1.1 Background

The standard assumption about economic agents is that they are exclusively self-interested. This is a useful theoretical benchmark. We know, in reality, that real human beings are also motivated by a range of other factors, including reciprocity, trust, emotions, and the well-being of others. These factors may shape economic outcomes in profound ways. An important question is therefore when, how, and why this happens, and in particular how such behavior interacts with policy interventions in various situations (including cooperation and health, for instance).

Cooperation is essential for the survival of our species and a fundamental aspect of human society (Axelrod and Hamilton, 1981; Bowles and Gintis, 2011; Fehr and Fischbacher, 2003; Gordon, 1954). Working together for a common good is an important part of our daily life. Examples include team efforts to deal with environmental issues, such as recycling and mitigation of climate change. It is therefore important to understand why and under what circumstances individuals cooperate. Two central aspects are the role of informal institutions and the design of formal institutions that can foster cooperation, on both a local and a global scale (Boyd et al., 2010; Fehr and Gächter, 2000; Fehr and Gächter, 2002; Hardin, 1968; Ostrom, 1990). This is the topic for the first two papers of the thesis.

One particularly important aspect in understanding human motivation is the influence of emotions and how they may shape economic outcomes (Elster, 1998; Loewenstein, 2000). Frustration and anger may drive revenge, punishment, and negative reciprocity. This has important strategic consequences and is the topic of the third paper of the thesis.

Health is another area of significant importance to human life, happiness and society at large. Medical costs are rapidly increasing and this poses a challenge for health care systems around the world and raises important questions regarding how to organize the system (Cutler, 2002; Newhouse, 1996). Since policy makers must rely on the behavior of physicians in implementing the overall established goals for the health care sector, a crucial issue is to understand how physicians' provision of health care responds to different types of reimbursement systems. This issue is investigated in the fourth paper of the thesis.

The main method used in all four papers consists of lab experiments. In the words of Vernon Smith (1976): "The laboratory becomes a place where real people earn real money for making real decisions about abstract claims that are just as 'real' as a share of General Motors" (p. 275). It is a (simple) microeconomic system that can link theory and controlled observation, and one can test central aspects of different institutional mechanisms (Smith, 1982). Two aspects are essential for this thesis: 1) Sharp theoretical predictions about human behavior can be tested and 2) the causal impact of institutions and incentive systems can be investigated in detail (Falk and Heckman, 2009).

1.2 Summary of the thesis

In **Paper I**, *Public Goods and Minimum Provision Levels: Does the Institutional Formation Affect Cooperation?* (with Peter Martinsson), we use a laboratory experiment to investigate how to overcome the problem of under-provision of public goods. Specifically, we examine the role of institutional formation on the implementation of a binding minimum contribution level in a linear public goods game. In the experiment, groups either face the minimum level exogenously imposed by a central authority or are allowed to decide for themselves by means of a group vote whether or not a minimum level should be implemented. We find a binding minimum contribution level to have a positive and substantial effect on cooperation. The main impact is on the extensive margin, meaning that it is possible to force free riders to increase their contribution without crowding out others' voluntary contributions. This result is robust to the mode of implementation, which is interesting from an institutional design point of view since it indicates that there are no positive effects from democracy per se in the context of our experiment. Taken together, our results show that when the minimum level is enforceable, it is a simple policy that will increase provision of the public good.

Paper II, *Framing and Minimum Levels in Public Good Provision* (with Peter Martinsson and Haileselassie Medhin), focuses on the design of institutions that foster

2

cooperation in the provision of local public goods, for example local environmental quality. Using a laboratory experiment in the field, we examine how the choice architecture of framing a social dilemma – *give to* or *take from* a public good – interacts with a policy intervention that enforces a minimum contribution level to the public good. We find that cooperation is significantly higher in the give frame than in the take frame in our standard public goods experiment. When a minimum contribution level is introduced, contributions are significantly higher in the take frame since contributions are crowded out in the give frame but crowded in in the take frame. Our results therefore stress the importance of choosing the frame when making policy recommendations. Moreover, the subjects in our experiments were farm household heads actively engaged in a number of public good-like decisions in real life, such as environmental rehabilitation and the maintenance of local infrastructure. Thus, this paper also contributes to extending the conventional analysis of cooperation to a non-student and non-Western subject pool, which is especially important since societies across the world differ substantially in terms of social organization, trust, fairness norms, and also in the nature of day-to-day cooperation problems.

Paper III, *Frustration and Anger in Games: A First Empirical Test of the Theory* (single authored), explores the strategic consequences of frustration and anger in human interaction. Angry individuals may become hostile in their dealings with others, and this emotional aspect could shape interaction and outcomes in, for example, situations involving negotiation and bargaining, contractual holdup, delegated decision making, conflict, and social dilemmas. Although it seems important to understand the sources of anger, as well as its consequences for strategic interaction, this topic has received relatively little attention in the development of behavioral theory. Battigalli, Dufwenberg, and Smith (2015) contribute to fill this gap in the literature by developing a formal framework where frustration and anger affect interaction and shape economic outcomes. In my paper, I design an experiment testing the predictions based on central concepts of their theory. The focus is on situations where other-responsibility is weak or nonexistent, and in this specific context I find only limited support for the theory: While unfulfilled expectations about material payoffs seem to generate negative emotions in subjects (which is in line with BDS' conceptualization of frustration), behavior is generally not affected by these emotions to the extent predicted by the theory.

Paper IV, *Physician Behavior and Conditional Altruism: The Effects of Payment System and Uncertain Health Benefit* (single authored), studies physician behavior using laboratory experiments. The focus is on the effect of the payment system (fee-for-service vs. capitation) on physicians' treatment decisions. More specifically, I investigate the altruistic behavior of physicians and whether this behavior is affected by

payment system and uncertainty in health outcome. The experiment shows that many physicians are altruistic toward their patients but also that the degree of altruism varies across patients with different medical needs, indicating that physicians condition their altruism on certain patient characteristics. Moreover, patients are overtreated in fee-for-service payment systems to the same extent as they are undertreated in capitation systems, and this result extends into domains of risk and uncertainty in patient health. There are more selfish physicians but also more purely altruistic physicians under fee-for-service than under capitation. Interestingly, even though the overall level of medical treatment differs substantially between capitation and fee-for-service payment, there is no significant difference in the distribution of physician types based on conditional altruism; the common categorization is that physician altruism is guided by severity of illness, both under capitation and fee-for-service.

References

Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211, 1390–1396.

Battigalli, P., Dufwenberg, M., Smith, A., 2015. Frustration and anger in games. IGIER Working Paper 539.

Bowles, S, Gintis, H., 2011. A Cooperative Species: Human Reciprocity and Its Evolution. Princeton University Press, Princeton

Boyd, R., Gintis, H., Bowles, S., 2010. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. Science 328, 617–620.

Cutler, D., 2002. Equality, efficiency, and market fundamentals: the dynamics of international medical-care reform. Journal of Economic Literature 40 (3), 881–906.

Elster, J., 1998. Emotions and economic theory. Journal of Economic Literature 36 (1), 47–74.

Falk, A., Heckman, J., 2009. Lab experiments are a major source of knowledge in the social sciences. Science 326, 535–538.

Fehr, E., Fischbacher, U., 2003. The nature of human altruism. Nature 425, 785–791.

Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. American Economic Review 90 (4), 980–994.

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

Gordon, H.S., 1954. The economic theory of a common-property resource: the fishery. Journal of Political Economy 62 (2), 124–142.

Hardin, G., 1968. The tragedy of the commons. Science 162, 1243–1248.

Loewenstein, G., 2000. Emotions in economic theory and economic behavior. American Economic Review 90 (2), 426–432.

Newhouse, J., 1992. Medical care costs: how much welfare loss? Journal of Economic Perspectives 6 (3), 3–21.

Ostrom, E., 1990. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press, Cambridge, UK.

Smith, V.L., 1976. Experimental economics: induced value theory. American Economic Review 66 (2), 274–279.

Smith, V.L., 1982. Microeconomic systems as an experimental science. American Economic Review 72 (5), 923–955.

Paper I

Public Goods and Minimum Provision Levels

Does the institutional formation affect cooperation?

Peter Martinsson[a], Emil Persson[a,*]

[a] *Department of Economics, University of Gothenburg, Sweden*

Abstract

We investigate the role of institutional formation on the implementation of a binding minimum contribution level in a linear public goods game. Groups either face the minimum level exogenously imposed by a central authority or are allowed to decide for themselves by means of a group vote whether or not a minimum level should be implemented. We find a binding minimum contribution level to have a positive and substantially significant effect on cooperation. The main impact is on the extensive margin, meaning that it is possible to force free riders to increase their contribution without crowding out others' voluntary contributions. This result is robust to the mode of implementation and thus when the minimum level is enforceable, it is a simple policy that will increase provision of the public good.

Keywords: Public goods, Minimum level, Voting, Experiment.
JEL Classification: C91, D72, H41.

---

* Corresponding author at: Department of Economics, University of Gothenburg, Box 640, SE 405 30 Gothenburg, Sweden; Tel.: +46 31 786 2927.
    E-mail addresses: emil.persson@economics.gu.se (Emil Persson), peter.martinsson@economics.gu.se (Peter Martinsson).

1. Introduction

Free riding is a frequently discussed reason for creating institutions to increase contributions to public goods. To reduce the problem of under-provision of public goods, formal institutions are often established. Many of these institutions focus specifically on the free riders or those who contribute a small amount to the public good for example by allowing punishment (e.g., see overview in Chaudhuri, 2011).[1] In real life, however, the most frequently used institution to combat free riders is to require some minimum contribution level. For example when it comes to the environment, there are certain standards for emissions from cars and for recycling at household level. Entrance fees to museums or national parks, and workload in teams, are other examples. The most common example is probably the financing of public goods such as health care and schools through taxation. The focus of this paper is whether a minimum level per se helps to increase contributions to a public good. Alternatively, these public goods could be provided in a regime without any minimum standards for cars or entrance fees, and rely solely on voluntary contributions. From an institutional design point of view, the question is whether a binding minimum level results in crowding out and whether the way the institution is implemented affects contribution levels. The objective of this paper is to investigate, by using a public goods experiment, the effect (i) of imposing a minimum contribution level for a public good and (ii) of the mode of implementation, i.e., whether the way the minimum level is introduced, either endogenously by majority voting or exogenously, affects contributions.

Binding minimum contribution levels have been studied experimentally with mixed results in terms of how levels of contributions to public goods have been affected. For instance, Andreoni (1993) and Gronberg et al. (2012) find that minimum levels increase contributions in public goods games with concave payoff functions and Chan et al. (2002) find the same effect but furthermore that it wears off as the minimum level is increased. Conversely, Kocher et al. (2016) implement minimum levels in a linear public goods game, and for a low minimum level they find no effect on public good provision whereas for a higher minimum level they find an overall positive effect. Moreover, the effects are qualitatively similar when the minimum level is endogenously imposed by one randomly selected group member. Keser et al. (2014) investigate minimum levels in a linear public goods experiment with endowment heterogeneity and find that the minimum level increases public good provision if subjects with a higher

---

[1] There are also experiments focusing on incentivizing people to give a high amount by allowing rewards (e.g., Sefton et al., 2007).

endowment face a higher minimum level.[2] Furthermore, a recent set of interesting papers use minimum levels to study strategic aspects of international negotiations concerning global public goods provision (e.g., climate protection). For example, Dannenberg (2012) and Dannenberg et al. (2014) study coalition formation in light of subsequent introduction of a minimum contribution level within this coalition, and Kesternich et al. (2014a,b) and Gallier et al. (2014) implement minimum contribution levels in settings where subjects are heterogeneous with respect to their endowment or their marginal per capita return from the public good.[3] Common among the latter set of papers is the feature that subjects simultaneously propose a minimum contribution level and the lowest level proposed is subsequently implemented.

When a minimum level is introduced, subjects who contribute below this level need to increase their contributions at least to the imposed level. But how do subjects who contribute to the public good above the minimum level even in the absence of such a mechanism react when the minimum level is implemented? These subjects could be *crowded out*, meaning that they voluntary decrease their contribution to the public good, e.g., from 50% to 40% of their endowment, or they could be *crowded in*, meaning that they voluntary increase their contribution to the public good, e.g., from 50% to 60% of their endowment. One reason for the latter effect could be that subjects are conditional cooperators, i.e., they want to cooperate if others cooperate and vice versa (e.g., Fischbacher et al., 2001) and therefore they increase their public good contribution since free riders are forced to increase their contribution at least up to the imposed minimum level. Several factors could result in crowding out. The implemented minimum level could serve as a focal point regarding which contribution level is expected. This could affect people in two ways: either directly by making them choose to contribute the minimum level or indirectly by making them believe that other group members will see the minimum level as a focal point. Crowding out could also be due to a "hidden cost of control," i.e., subjects do not like the control exerted upon them and therefore reduce their contribution. Falk and Kosfeld (2006) and Ziegelmeyer et al. (2012) implement principal-agent experiments where agents receive a fixed wage and can choose to engage in a production activity that is costly to the agent but beneficial to the principal; the "hidden cost of control" amounts to the degree to which agents reduce their efforts when the principal implements a minimum performance requirement. The effect is strong in the former study but somewhat weaker in the latter (where it is dominated by

---

2 For public goods experiments using non-binding minimum levels, see, e.g., Galbiati and Vertova (2008) and Riedel and Schildberg-Hörisch (2013).
3 A related paper is Gerber et al. (2013), who examine minimum requirements with respect to formation of the coalition itself.

the positive effect of forcing low-performing agents to increase their effort). This effect is potentially magnified by subjects' beliefs about others' reactions following the introduction of the minimum contribution level.

If hidden costs of control are pervasive, there is a risk that they will partly or even completely offset the direct effect of the minimum contribution level (forcing free riders to increase their contribution). In this case, legitimizing the minimum level by granting local decision rights to groups may be critical in order to achieve a more successful outcome. We test this by allowing groups to decide for themselves whether or not the minimum level should be implemented. In more detail, we study experimentally the implementation of a minimum contribution level in a linear public goods game.[4] Groups either face the minimum level exogenously imposed by a central authority or are allowed to choose whether or not to implement it based on the outcome of majority voting.

A growing experimental literature examines how centralized institutions with exogenous power to implement policies can increase cooperation in social-dilemma situations. In this setting, the degree of crowding out due to hidden costs of control and whether it can be alleviated through endogenous institutional formation is still an open empirical question, which we address in the current paper. We find a minimum contribution level to have a positive and substantial effect on cooperation. The main impact is on the extensive margin, meaning that it is possible to force free riders to increase their contribution without crowding out others' voluntary contributions. This result is robust to the mode of implementation; that is, exogenous minimum levels are as effective as endogenous minimum levels. This result stands in contrast to several recent papers documenting that endogenously chosen institutions do alleviate collective-action problems to a larger extent than when the same institutions are exogenously imposed. Tyran and Feld (2006) let subjects vote on the introduction of a law that prescribes full contribution to the public good. They find that a "mild law," which is backed up by a non-deterrent sanction, is more effective when chosen endogenously than when exogenously imposed. Dal Bó et al. (2010) introduce the possibility for

---

[4] The main difference between our design using a linear payoff function and other experimental studies of minimum levels that instead use concave payoff functions (e.g., Andreoni, 1993; Chan et al., 2002; Gronberg et al., 2012) is that the minimum level is set above rather than below the Nash equilibrium. This is important since we want to consider the issue of conditional cooperation in combination with hidden costs of control and these potential effects become more salient with linear payoff functions. Papers using concave payoff functions have studied a different question, namely the neutrality of government spending with respect to privately provided public goods (Warr, 1982; Bergstrom et al., 1986; Andreoni, 1989, 1990).

subjects to endogenously impose a fine on unilateral defection in a prisoners' dilemma game, and find that this increases cooperation compared with the same fine exogenously imposed. Sutter et al. (2010), Markussen et al. (2014a), and Kamei et al. (2015) find that peer-to-peer punishment can sustain high levels of cooperation to a greater extent when democratically chosen than when exogenously imposed. However, Markussen et al. (2014b) find no effect of endogenous institutional choice. They investigate the effect of intergroup competition on local public good provision within groups, and find that competition is as effective if endogenously chosen as if exogenously imposed. Likewise, Sutter and Weck-Hannemann (2004) extend Andreoni (1993) by introducing endogenous taxation on public good contributions and find no positive effect on contributions compared with exogenous taxation. Taken together, the benefit of granting groups participation rights seems to be context dependent and we discuss our results in light of this in a concluding section of the paper.

A majority of the groups in the endogenous treatment in our experiment successfully implement the minimum-level institution. This is in line with several other papers that examine endogenous institutional choice and find that groups – when given the freedom of choice – successfully implement efficiency-enhancing institutions. In Kroll et al. (2007), subjects submit proposals for how much everyone should contribute to the public good. Then they vote on the proposals and if one proposal receives sufficient support it is implemented. They find that when the vote is binding, it results in high levels of public good provision. Putterman et al. (2011) let subjects vote on the parameters of a centralized punishment scheme in the context of public good provision and find that most groups are able to implement a scheme where it is optimal for everyone to contribute their entire endowment. In Markussen et al. (2014a), groups endogenously choose between different types of punishment schemes. They find that most groups are able to implement low cost and deterrent sanctions, which can increase public good provision. Kosfeld et al. (2009) implement public goods experiments where groups can form an organization in which all members must contribute their entire endowment to the public good, while subjects in the group not covered by the organization can contribute as they wish. They find that contributions and welfare are higher when groups are allowed to form organizations. Finally, in Ertan et al. (2009), subjects endogenously choose the rules for peer-to-peer punishment. They find that most groups eventually implement a scheme that allows them to achieve high levels of public good provision.[5]

---

[5] See also Dal Bó (2014) for an overview of the experimental literature on collective action and the

2. Experimental design

We use a linear public goods experiment with a design similar to the one developed by Fischbacher et al. (2001) and where the main characteristic is that the strategy method is used. Each group consists of three randomly matched members, and the marginal per capita return (MPCR) from the public good is 0.5. For a rational and selfish subject, an MPCR below one results in a dominant strategy to free ride, i.e., to contribute zero to the public good. However, since MPCR*$n > 1$, where n is number of group members, it is socially optimal to contribute the whole endowment. Thus, our choice of marginal per capita return creates the conflict between private and social optima that characterizes a public good. Each subject has an endowment of 20 tokens, and we denote c the amount invested in the public good. Thus, the payoff for subject i is given by

$$\pi_i = 20 - c_i + 0.5 \sum_{j=1}^{3} c_j. \tag{1}$$

The key difference between the Fischbacher et al. design and standard public goods experiments is that the former employs the strategy method. The subjects make two different types of decisions on how much they would like to contribute to the public good – one unconditional and one conditional. In the unconditional decision, the subject states how much she would like to contribute to the public good as in a standard public goods experiment. The conditional contribution to the public good is the amount a subject would like to contribute conditional on the other two group members' average contribution. Each subject makes 21 conditional decisions in total, one for each integer number from 0 to 20, representing all the possible (integer) averages of the other two group members. All contribution decisions are incentive compatible since the payoff-relevant decision for two randomly selected members of the group is the unconditional contribution, and by using their average unconditional contribution, the contribution of the third member is determined as the conditional contribution for that specific average contribution.

The experiment consists of two stages, both of which are based on the design by Fischbacher et al. (2001): (i) a standard public goods experiment and (ii) a public goods experiment where a minimum contribution level is either exogenously or endogenously imposed. There are two treatments in the experiments: an exogenously imposed minimum contribution level and an endogenously imposed minimum contribution level. In the exogenous treatment, a minimum contribution level of 5 tokens is imposed, while in the endogenous treatment the implementation of a minimum level of 5 tokens is

---

effects of democratic institutions.

subject to a group vote. There is no cost of voting, and the voting is made mandatory. A majority rule determines the outcome, and thus if at least 2 out of the 3 group members vote for the minimum contribution level of 5 tokens it is imposed, and vice versa.

To test for the impact of a minimum contribution level, we run two types of sessions. Both of them begin with the baseline game described in the previous paragraph (and without any minimum level). In sessions with the exogenous treatment, the standard public goods experiment is followed by the same game but now exogenously imposing the rule that everyone must contribute at least 5 tokens to the public good. In sessions with the endogenous treatment, the standard public goods experiment is instead followed by a voting stage where each member casts a vote on whether or not they would like to introduce a minimum level of 5 tokens. Thus, the majority outcome of the vote determines whether or not a minimum level is imposed. This means that the baseline game is followed by either the standard public goods experiment again or the same experiment but with a minimum level of 5 tokens depending on the majority decision in the voting stage. We use perfect stranger matching and the subjects are not informed about the outcome of the first public goods experiment until the final payment is made. This is clearly stated in the instructions given to the subjects (see Appendix).[6] Thus, we will be able to document the importance of participation rights if the endogenous minimum level either increases contributions to a larger extent or decreases contributions to a lesser extent than does the exogenous minimum level.

The logistics of the experiment were as follows. The standard public goods experiment was first explained and then conducted. The experiment instructions included examples and individual exercises to be solved to ensure that subjects understood the setup. Upon completion of the exercises, they were also solved in public by the experimenter in order to facilitate understanding. Questions were answered in private throughout the experiments. The subjects were randomized into a session with either exogenous or endogenous imposition of the minimum level. The voting stage preceded the second public goods game in sessions with the endogenous treatment, where subjects were asked to vote either for or against imposition of a minimum contribution level of 5 tokens. The exogenous treatment was identical except for the voting, and thus a paragraph was excluded from the instructions, which otherwise were exactly the same.

---

[6] We do not test for order effects. The reason is that the imposed order is natural, in the sense that we are interested in the effect of a minimum-level policy when subjects are experienced with the baseline (voluntary contribution mechanism). This is what typically happens in reality: a collective-action problem is identified and a remedy – the minimum level – is sought and implemented. Differences across treatments are tested as the difference between within-subject estimates.

After both experiments were finished, we elicited each subject's belief about others' contributions; subjects were asked to guess the average contribution of the other group members, and correct guesses were monetarily rewarded. Note that the outcome of the public goods experiment was not revealed to the subjects until the end of the experiment. Finally, subjects answered a post-experiment questionnaire related to socio-economic background factors. The experiment was conducted in a total of six sessions at the University of Gothenburg using pen and paper. Subjects were recruited using ORSEE (Greiner, 2015). An experimental session lasted on average 90 minutes, and one token corresponded to 2 SEK (at the time of the experiment corresponding to 0.30 USD). A total of 144 subjects participated, 51 in the exogenous treatment and 93 in the endogenous treatment,[7] each earning an average of 150 SEK (22.50 USD).

3. Results

We begin the analyses by focusing on the unconditional contributions, which are summarized in Table 1. In the standard public goods experiment, the averages of the unconditional contributions are very similar among the subjects who in the second stage were allocated to either the exogenous or the endogenous treatment, i.e., 8.35 and 8.24 tokens, respectively. The small difference in contributions is insignificant, indicating that the randomization of subjects between treatments was successful (Mann-Whitney U test; $p = 0.93$).[8] The contribution levels in the standard public goods experiment correspond to slightly more than 40% of the endowment, which is in line with other public goods experiments (e.g., Chaudhuri, 2011). As expected, contribution levels are higher in the public goods experiment with minimum levels (MCL). When the minimum level is exogenously imposed, the average contribution goes up to 9.43, implying a statistically significant increase of 1.08 tokens (Wilcoxon signed-rank test; $p = 0.01$). In the endogenous treatment, the corresponding effect is a 0.85-unit increase in average contribution and again the effect is significant (Wilcoxon signed-rank test; $p < 0.01$).

In the endogenous treatment, the group members voted about imposing a minimum contribution level of 5 tokens. A majority of the subjects voted in favor of the minimum level. Of the 93 subjects in the endogenous treatment, 75 (81%) voted in favor of the minimum contribution level and 18 voted against it. Correspondingly, 27 of the 31

---

[7] We choose a larger sample size in the endogenous treatment to avoid too few groups voting in favor of imposing the minimum level.

[8] We report from two-sided tests throughout the paper.

groups (87%) voted for imposing a minimum contribution level. Interestingly, we find that men are much more likely to vote no. Fourteen of the 18 no-voters (78%) are men; in total, 28% of the men but only 10% of the women oppose the minimum-level institution by vote. The difference in proportions is significant at the 5% level (Kruskal-Wallis test; $p = 0.03$).

The attractive feature of a minimum contribution level is that it forces any free riders, or low contributors, to increase their contribution up to the imposed mandatory minimum. In addition to this effect, subjects contributing more than the imposed minimum can voluntarily adjust their contribution: they are crowded in if they voluntarily increase their contribution and crowded out if they voluntarily decrease their contribution. For example, subjects could be crowded out since the minimum level can signal distrust similar to the hidden cost of control in Falk and Kosfeld (2006), or it may serve as a focal point. Conditional cooperators can also be crowded in, since the minimum level forces free riders to increase their contribution, or they can be crowded out due to the belief that others will react negatively to the hidden costs of control or due to the belief that the minimum level will serve as a focal point. Thus, the overall effect is ex-ante unclear and we analyze the effect of the minimum level conditional on the contributions in the baseline case.

The main behavioral difference between the exogenous and endogenous treatments is that the hidden costs of control can be avoided in the latter case since subjects can decide for themselves on institution formation rather than being forced to contribute the minimum level. That is, when groups are given the freedom of choice, do they on average increase their public good contribution *to a larger extent* than groups not given this opportunity? We investigate this by comparing the impact of a minimum provision level between the exogenous and endogenous treatments. We cannot reject the null hypothesis of no effect of endogenous institutional choice; the difference (in differences) between the treatments is 0.04 tokens in public good contributions (Mann-Whitney U test; $p = 0.57$). Thus, the feature of having the possibility to choose institution has a zero net effect on contributions. This seems to indicate that the main driver of the result is the direct effect of increased contributions from free riders and low contributors, which we analyze in more detail below.

Result 1. *Minimum contribution levels have a significantly positive effect on public good contributions.*

Result 2. *A majority of the groups implemented the minimum level in the endogenous treatment.*

Result 3. *The endogenous and exogenous treatments have similar significant positive effects on public good contributions, indicating a low net effect of hidden costs of control.*

Table 1 about here

Table 2 shows contributions and beliefs in the two treatments. In both the exogenous and the endogenous treatment, subjects increase their belief about others' contribution when the minimum level is implemented (Wilcoxon signed-rank tests; $p < 0.01$ in either case). One advantage in the endogenous treatment is that subjects are free to decide whether or not to impose the minimum level. If the hidden cost of control is expected to be a substantial problem in the case of exogenous minimum levels, there are mainly two positive effects of endogenous formation of (minimum-level) institutions. First, if individuals value the freedom of choice per se, the hidden cost of control might be overall reduced when groups are granted the right to decide for themselves whether or not to introduce minimum levels. Second, some individuals might be more control averse than others. Thus, by allowing for endogenous formation of minimum-level institutions, those groups where the exogenous minimum level would be most likely to crowd out voluntary contributions are allowed to endogenously opt out, by voting against the implementation.

The bottom panel of Table 2 shows the changes in contributions and beliefs by voting and group decision in the endogenous treatment. By and large, subjects should vote yes unless they believe that contributions were already at substantial levels ("high beliefs" hereafter) in the baseline (or if they are risk averse and uncertain about the accuracy of their high beliefs) or they have a strong aversion against control.[9] A yes-vote signals that individuals are (i) selfish and believe that others contribute at low levels (the minimum level is a Pareto improvement), (ii) conditional cooperators who believe that others contribute at low levels (the minimum level enforces conditional cooperation at higher contribution levels), and/or (iii) not strongly averse to control. Thus, in a Yes-Yes group ("yes" for own and group vote, respectively), we expect an increase in both

---

[9] To a selfish player with high beliefs, the minimum level will only have a direct effect on own contributions (which she will be forced to increase), and, furthermore, a yes-vote does not work as a "bait strategy" since conditional cooperators are already contributing substantially (according to her belief). To a conditional cooperator with high beliefs, the minimum level will not have a strong direct effect on others' contributions since according to her belief, nobody is free riding. Furthermore, a subject with high beliefs about others' contributions might be more likely to fear that others will be crowded out due to the hidden costs of control (and therefore vote no). However, if subjects are uncertain about the accuracy of their beliefs, risk-averse individuals might still vote yes in order to hedge against the possibility that their high beliefs are wrong.

contributions and beliefs about others' contributions, since subjects who vote yes are either selfish or conditional cooperators and either type would respond by an increase in contributions and believe that others in the group do the same. This is what we see in Table 2: For the 72 individuals who voted yes and ended up in a group where the majority voted yes, contributions and beliefs significantly increase following the implementation of the minimum level (Wilcoxon signed-rank tests; $p < 0.01$ in either case). No-voters who nevertheless end up in a group where the majority voted yes (No-Yes) face the same restriction ex-post, i.e., the minimum level. For these individuals, the impact on own contributions should be negative if the individual voted no due to a strong aversion against control, but not if the no-vote was due to a fear that others in the group were averse toward control, since they seem not to be since they voted yes. The predicted change in contribution is thus ambiguous for this group, and we cannot reject a null hypothesis of no change when the minimum level is implemented (Wilcoxon signed-rank test; $p = 0.63$). In neither of the two bottom rows of the table (No-No and Yes-No) do we expect much of a change in contributions or beliefs since the minimum level is not implemented. This is also in line with the results in the table.[10]

We have argued that the hidden cost of control is the main reason for voting against the minimum level, either directly or indirectly due to a fear that others will react negatively, and find that the results in Table 2 point in this direction since subjects in Yes-Yes groups on average increase their contributions under minimum levels more than subjects in No-Yes groups. Interestingly, No-Yes groups do not increase their contributions more than No-No groups even though in the former, free riders and low contributors are forced to contribute at least at the minimum level.

Table 2 about here

Result 4. *When a group majority decides to implement the minimum contribution level, subjects voting yes increase their contributions significantly more than subjects voting no, while there is no significant difference in contribution behavior between the two groups of voters in the standard public goods experiment.*

Subjects could be crowded out or crowded in by the minimum level and here we

---

[10] In No-No groups, the minimum level is not implemented and this is in line with the own vote (the subject got what she wanted). In Yes-No groups, if the individual is a pessimistic conditional cooperator (and hence voted yes), there is no expected increase in others' contributions to respond to, and if the individual was selfish and voted yes because of an expected increase in everyone's contributions, we would not expect a behavioral response regardless (we would only expect a forced increase in contributions if the minimum level were to be implemented).

investigate this issue in more detail. We plot individual contributions before and after the minimum provision level is imposed in the exogenous (Figure 1) and endogenous (Figure 2) treatments, respectively.[11] In the figures, on the x-axis we show contribution in the standard public goods game, while on the y-axis we show the contribution when the minimum level is imposed. Any observation below the 45-degree line shows a higher contribution in the baseline than when the minimum level is imposed, i.e., crowding out, whereas any observation above the 45-degree line shows crowding in for subjects who contributed at least 5 tokens in the baseline, i.e., the imposed minimum level. The figures indicate substantial individual heterogeneity as to whether subjects are crowded in, crowded out, or unaffected by the minimum level.

Figures 1 and 2 about here

Tables 3–5 add more detailed information. Table 3 shows that 49% and 43% increased their contribution under the minimum level in the exogenous and endogenous treatment, respectively, and this corresponds to the number of observations above the 45-degree line in Figures 1 and 2. Conversely, 20% and 17% decreased their contribution in the respective treatment. The positive effect clearly dominates, and we know this from Result 1, but the values include the direct effect of forcing free riders and low contributors to contribute at least the enforced minimum. Therefore, in Tables 4 and 5 we separate out this direct effect (top row) from subjects' behavioral responses to the minimum level, for the exogenous and endogenous treatment, respectively. In the exogenous treatment (Table 4), 23% of the subjects contributed nothing in the standard public goods experiment and exactly 5 tokens when this was the enforced minimum, and 8% increased their contribution from 5 to an average of 8.25 tokens under the minimum level, indicating that they were crowded in. Conversely, 12% of subjects reduced their contribution from an average contribution of 8.83 in the baseline all the way down to the enforced minimum level, and they were thus crowded out by the minimum level. The average contribution of the remaining 49% was virtually unchanged. The pattern is similar for the endogenous treatment (Table 5), with the main difference that crowding out is stronger: 10% of the subjects drop from an average of 11.38 all the way down to the enforced minimum of 5 tokens. By and large, crowding in seems to occur mainly at low levels of standard public goods contributions, which perhaps was caused by pessimistic conditional contributors, whereas crowding out is more likely for subjects who already contribute substantially to the public good.

---

[11] From here on we drop the four groups in the endogenous treatment in which the majority voted against the minimum level.

However, for many subjects, the average contribution is rather stable and thus the main share of the positive overall minimum-level effect demonstrated in Table 1 comes from the direct effect of pushing free riders and low contributors up to the minimum level. For behavioral responses beyond this, the net effect is more or less zero on average.

Tables 3, 4, and 5 about here

Result 5. *Minimum contribution levels effectively work on the extensive margin; there is no evidence that it will serve to crowd out voluntary contributions to the public good.*

Previous empirical research has suggested that a large fraction of people are conditional cooperators, i.e., they cooperate if others cooperate and vice versa (e.g., Fischbacher et al., 2001). The unconditional results in Table 2 indicate a correlation between contributions and beliefs. In Figures 3 and 4, we explore how contributions match beliefs about others' contributions. The difference in contributions between the MCL public goods experiment and the standard public goods experiment (y-axis) is plotted against the corresponding difference in beliefs about others' contributions (x-axis). On average, individuals who believe that the aggregate public good provision will increase also increase their own contribution. For those who instead believe in a reduction of contributions, the opposite is true, i.e., they decrease their own contribution. Behavior in our experiment is thus consistent with the notion of conditional cooperation.

Figures 3 and 4 about here

The other part of the public goods experiment elicited conditional contributions.[12] Conditional contributions are elicited using a variant of the strategy method. The main advantage with this procedure is that we can control for subjects' beliefs about others' contributions in an incentive-compatible way using the strategy method and furthermore classify subjects into different contributor types. We can thus in more detail analyze the direct impact of the policy on subjects' cooperative preferences. For instance, those who in the baseline contribute substantially to the public good regardless of others' contributions might shift their schedule downwards due to hidden costs of control; or conditional cooperators might shift their schedule upwards at low levels of others' average contribution if they like the idea that everybody has to contribute to the public

---

[12] One subject in the endogenous treatment contributed less than the required minimum of 5 tokens in the MCL contribution table. Since all entries were below the minimum level, we treat this subject in our analyses as if she had contributed 5 tokens throughout the contribution table.

good under the new institution.[13]

Figures 5 and 6 show the average conditional contribution for each level of average contribution level of the others separated by treatment, i.e., endogenous and exogenous minimum level, respectively. The average own contributions to the public good (y-axis) are plotted against the conditional average contributions of the other two group members (x-axis), following Fischbacher et al. (2001). As expected, there is an overall positive relationship between own contributions and others' average contributions. These results show that subjects are on average conditional contributors since their own contribution levels increase when the average contributions by others increase. Interestingly, under the minimum contribution level, the slopes of the schedules are approximately the same as for the standard public goods experiment, but there is a shift in levels for a given average contribution by others. This corroborates the finding that minimum provision levels are effective; free riders can be forced to increase their contribution without ex-ante compliers reacting against the control exerted upon the group as a whole. The results hold if beliefs about others' average contributions remain unchanged or do not decrease more than 2–3 tokens.

Figures 5 and 6 about here

In Table 6, we follow up more closely on the individual data used in Figures 5 and 6. We regress own conditional contributions on others' average contributions. In column 1, we pool the sample and investigate the impact of the minimum level on the intercept and slope of the regression line, where the latter measures the marginal increase in contribution when others' average contribution increases with one token. The MCL dummy captures the impact of the minimum level. For instance, when the average contribution of the other group members equals 5 tokens, subjects on average report an own contribution of 3.62 ($1.37 + 5{\times}0.45$) tokens in the baseline and this number increases to 5.82 ($1.37 + 5{\times}0.45{+}2.57 - 5{\times}0.07$), i.e., by 2.22 tokens, following the implementation of the minimum level. In column 2, we investigate the full within-between dimension of the data, comparing differences between the endogenous and exogenous treatments in the effect of the minimum level on conditional contributions. Neither with respect to intercept (*Endo $\times$ MCL*) nor slope (*Endo $\times$ MCL $\times$ Others*) is the MCL effect different between the treatments, since the estimated effects are both

---

small and insignificant. Thus, there is no evidence of a more pronounced minimum-level effect in the endogenous treatment vis-à-vis the exogenous treatment, which is in line with previously reported descriptive statistics.

Table 6 about here

In more detailed analyses, we categorize subjects into contributor types following the convention introduced by Fischbacher et al. (2001). They classified subjects into four types: conditional cooperators, free riders, hump-shaped contributors, and others.[14] Tables 7 and 8 show the distribution of types both in baseline and with the minimum contribution level, for the exogenous and endogenous treatment, respectively. The tables enable us to investigate type stability, which is related to the discussion about crowding in and crowding out as a behavioral response to the minimum-level institution, and again we can analyze this issue while controlling for beliefs about others' behavior.

Most notably, in both treatments, the share of conditional cooperators decreases whereas the share of free riders increases. For example, in the exogenous treatment (Table 7), the share of free riders increases from 17.65% in the baseline to 25.49% under the minimum level; 5.88 of this 7.84 percentage point increase constitute baseline conditional cooperators. Of the 64.7% baseline conditional cooperators, only 52.94 percentage points remain under the minimum level. The pattern is very similar in the endogenous treatment. To fully account for behavioral responses, we also present, in Table 9, the type distribution when the types in the baseline are characterized *after* we manually adjust all subjects up to the minimum of 5 tokens in the standard public goods experiment. Thus, any remaining effect comes from subjects' behavioral reaction to the implementation of the minimum level. Interestingly, when we characterize baseline types in relation to the adjusted minimum level, contributor types are remarkably stable across baseline and the minimum-level institution. For example, when a baseline free rider is characterized as someone who never contributes above 5 tokens (rather than someone who never contributes above 0 tokens), there are 25.49% free riders already in the exogenous baseline and thus there is no aggregate increase in the share of free riders when we measure types in this way. In the endogenous treatment, there is a 3.70

---

[14] Based on their conditional contribution schedule, we define types as follows: Conditional cooperators have either a weakly increasing schedule or the Spearman correlation (with the others' average contribution) is positive and significant at the 1% level; free riders always contribute the lowest amount allowed; hump-shaped contributors display a positive (negative) Spearman correlation up to (beyond) their highest contribution, both significant at the 1% level; and "others" do not fit into any of these categories. For a discussion on classification, see, e.g., Fischbacher et al. (2001).

percentage point increase in the share of free riders, but this is a substantially smaller rise than the 16.05 percentage point increase we observe in Table 8 when we use the normal type characterization without the adjustment of a minimum level of 5 tokens in the standard public goods experiment. Thus, the change in type distribution brought about by the minimum level seems to be an artefact of the minimum level automatically classifying weak contributors as free riders.[15] In conjunction with the regression results in Table 6, the absence of treatment-driven differences in the distribution of contribution types supports our main conclusion regarding the impact of endogenous institution formation on cooperative preferences.

Tables 7, 8 and 9 about here

Result 6. *Allowing groups to decide for themselves whether or not a minimum provision level should be implemented does not make individuals more cooperative. Vis-à-vis exogenous implementation, the opportunity to vote does not affect subjects' cooperative preferences.*

4. Discussion and conclusion

In this paper, we investigate the effect of imposing a minimum contribution level on cooperative behavior. A centralized institution with exogenous imposition of rules and regulations of minimum levels is in many cases the only feasible option for policy makers since letting individuals themselves be responsible for punishment, rewards, and potentially ostracizing others is not possible in many situations, including taxation and environment. Without considering behavioral responses, we would expect an imposed minimum contribution level to increase cooperation levels if the society consists of free riders or subjects who contribute a small amount. However, imposing a minimum level might add an additional control cost as discussed by Falk and Kosfeld (2006), which could result in reduced contributions to the public good among subjects who already contributed above the imposed minimum level. Previous research has also shown that it is not only the introduction of an institution per se that might affect behavior, but also the way it is introduced (e.g., Dal Bó et al., 2010; Markussen et al., 2014a; Sutter et al., 2010; Tyran and Feld, 2006).

The objective of this paper was to investigate, by using a public goods experiment, the effect of (i) imposing a minimum level of contribution to a public good and (ii) the

---

[15] For instance, a conditional contributor with a positive contribution schedule that never reaches above 5 tokens would be classified as a free rider in Table 9 once the minimum level has been implemented.

mode of implementation, i.e., whether the way in which the minimum level is introduced, either endogenously by majority voting or exogenously, affects contributions.

We believe that a low minimum level, which we in our experiment set to 25% of the endowment, is the most feasible policy option since too high exogenous minimum levels would be met with resistance since people differ in terms of their preferences, their endowments, and in some cases also in their costs of contributing to the public good, e.g., costs of effort. In addition, too high endogenous minimum levels would likely be opposed and voted down for the same reasons. An alternative could be to introduce non-binding minimum levels, but empirical evidence indicates that non-binding announcements and promises have little or no effect on voluntary contributions to public goods (e.g., Bochet and Putterman, 2009).

Overall, our results show that a minimum contribution level has a positive and sizeable effect on cooperation whether imposed exogenously or endogenously. This positive effect is mainly due to the fact that free riders are forced to increase their contribution, while at the same time the contributions by those who originally contributed above the imposed minimum level were not crowded out. In the context of cooperation, there does not seem to be a hidden cost of control. This is probably linked to the fact that imposing a minimum level actually has a potentially positive welfare effect per se, while in the case of Falk and Kosfeld (2006) the control mechanism is directly linked to distrust. However, we cannot rule out a spillover effect of the imposed minimum level in one area on other areas with clearer elements of trust, which would then be an indirect negative welfare effect of the policy.

Our findings are robust to whether the minimum contribution level is imposed exogenously or endogenously by voting. Thus, the overall effect of majority voting as a signal of cooperation does not seem to be strong in this context. An explanation is that there is an incentive for selfish players to vote in favor of the minimum level, since it has a potentially positive welfare effect per se, which makes it difficult for intrinsically cooperative individuals to signal their cooperative intentions by voting yes. By comparison, in Tyran and Feld (2006), voting for a non-deterrent law of full contribution is a strong signal of cooperative intentions since selfish players would vote no unless they think that there are conditional cooperators in the group who would take a yes vote as a signal of cooperative intentions. Likewise, in Dal Bó et al. (2010), the only reason a selfish player would vote to impose a fine on unilateral defection in a prisoners' dilemma game would be to use it as a "bait strategy," i.e., to signal that one intends to cooperate when in reality one plans to defect (since the payoff from defection

is higher if the other player cooperates). Similarly, in Sutter et al. (2010) and Markussen et al. (2014a), subjects voting in favor of introducing peer-to-peer punishment signal their intention to make use of this mechanism in order to enforce higher levels of cooperation. To summarize, the main difference between our paper and for instance the four papers discussed above is that even if nobody wants to cooperate, there is still an incentive to vote yes in our setting. This is not the case in any of the other papers.

Still, there could be reasons beyond signaling for why individuals cooperate more under endogenous institutions. Dal Bó et al. (2010) discuss the establishment of a cooperative norm and coordination as possible reasons beyond signaling, and furthermore show that signaling is not the main driver behind a sizeable effect of democratically chosen institutions on cooperation in their setting. In our setting, one such effect could be that endogenous formation of institutions reduces the hidden costs of control and thus has a legitimizing effect on ex-post cooperation. The likely explanation for an absence of such an effect is that a hidden cost of control seems not to be pervasive in the context of cooperation. The societal welfare effect of a minimum-level policy depends on a number of factors, including the perceived control cost and the signaling effect of voting in favor of imposing the minimum level. An important ingredient for the welfare effect is the composition of contributor types, and how the above mentioned policy variables affect this composition. Societies with a high fraction of free riders or where there is a strong and positive signal of increased contributions when imposing a minimum level will clearly benefit from the policy, while the opposite might hold for societies with high trust levels and a large fraction of conditional cooperators. Since minimum levels are used in many areas of society, it is important to understand their welfare effects, especially considering that the imposed minimum level often is low.

Acknowledgement

References

Andreoni, J., 1989. Giving with impure altruism: applications to charity and Ricardian equivalence. Journal of Political Economy 97 (6), 1447–1458.

Andreoni, J., 1990. Impure altruism and donations to public goods: a theory of warm-glow giving. Economic Journal 100 (401), 464–477.

Andreoni, J., 1993. An experimental test of the public-goods crowding-out hypothesis. American Economic Review 83 (5), 1317–1327.

Bergstrom, T., Blume, L., Varian, H., 1986. On the private provision of public goods. Journal of Public Economics 29 (1), 25–49.

Bochet, O., Putterman, L., 2009. Not just babble: opening the black box of communication in a voluntary contribution experiment. European Economic Review 53 (3), 309–326.

Chan, K.S., Godby, R., Mestelman, S., Muller, R.A., 2002. Crowding-out voluntary contributions to public goods. Journal of Economic Behavior and Organization 48 (3), 305–317.

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

Cubitt, R.P., Drouvelis, M., Gächter, S., Kabalin, R., 2011. Moral judgments in social dilemmas: how bad is free riding? Journal of Public Economics 95 (3–4), 253–264.

Dal Bó, P., 2014. Experimental evidence on the workings of democratic institutions, in: Galiani, S., Sened, I. (Eds.), Institutions, Property Rights, and Economic Growth: The Legacy of Douglass North, Cambridge University Press, pp. 266–288.

Dal Bó, P., Foster, A., Putterman, L., 2010. Institutions and behavior: experimental evidence on the effects of democracy. American Economic Review 100 (5), 2205–2229.

Dannenberg, A., 2012. Coalition formation and voting in public goods games. Strategic Behavior and the Environment 2 (1), 83–105.

Dannenberg, A., Lange, A., Sturm, B., 2014. Participation and commitment in voluntary coalitions to provide public goods. Economica 81 (322), 257–275.

Ertan, A., Page, T., Putterman, L., 2009. Who to punish? Individual decisions and majority rule in mitigating the free rider problem. European Economic Review 53, 495–

511.

Falk, A., Kosfeld, M., 2006. The hidden costs of control. American Economic Review 96 (5), 1611–1630.

Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. American Economic Review 90 (4), 980–994.

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

Galbiati, R., Vertova, P., 2008. Obligations and cooperative behavior in public good games. Games and Economic Behavior 64 (1), 146–170.

Gallier, C., Kesternich, M., Sturm, B., 2014. Voting for burden sharing rules in public goods games. ZEW Discussion Paper 14-056.

Gerber, A., Neitzel, J., Wichardt, P., 2013. Minimum participation rules for the provision of public goods. European Economic Review 64, 209–222.

Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. Journal of the Economic Science Association 1 (1), 114–125.

Gronberg, T.J., Luccasen III, R.A., Turocy, T.L., Van Huyck, J.B., 2012. Are tax-financed contributions to a public good completely crowded-out? Experimental evidence. Journal of Public Economics 96 (7–8), 596–603.

Henrich, J., Heine, S.J., Norenzayan, A., 2010. The weirdest people in the world? Behavioral and Brain Sciences 33 (2–3), 61–135.

Kamei, K., Putterman, L., Tyran, J.-R., 2015. State or nature? Endogenous formal versus informal sanctions in the voluntary provision of public goods. Experimental Economics 18, 38–65.

Keser, C., Markstädter, A., Schmidt, M., 2014. Mandatory minimum contributions, heterogeneous endowments and voluntary public-good provision. CIRANO Scientific Series 2014s-47.

Kesternich, M., Lange, A., Sturm, B., 2014a. On the performance of rule-based contribution schemes under endowment heterogeneity. ZEW Discussion Paper 14-055.

Kesternich, M., Lange, A., Sturm, B., 2014b. The impact of burden sharing rules on the voluntary provision of public goods. Journal of Economic Behavior and Organization

105, 107–123.

Kocher, M, Martinsson, P., Persson, E., Wang, X., 2016. Is there a hidden cost of imposing a minimum contribution level for public good contributions? Working paper.

Kosfeld, M., Okada, A., Riedl, A., 2009. Institution formation in public goods games. American Economic Review 99 (4), 1335–1355.

Kroll, S., Cherry, T.L., Shogren, J.F., 2007. Voting, punishment, and public goods. Economic Inquiry 45 (3), 557–570.

Markussen, T., Putterman, L., Tyran, J.R., 2014a. Self-organization for collective action: an experimental study of voting on sanction regimes. Review of Economic Studies 81 (1), 301–324.

Markussen, T., Reuben, E., Tyran, J.R., 2014b. Competition, cooperation and collective choice. Economic Journal 124 (574), 163–195.

Putterman, L., Tyran, J.R., Kamei, K., 2011. Public goods and voting on formal sanction schemes. Journal of Public Economics 95 (9–10), 1213–1222.

Riedel, N., Schildberg-Hörisch, H., 2013. Asymmetric obligations. Journal of Economic Psychology 35, 67–80.

Sefton, M., Shupp, R., Walker, J., 2007. The effect of rewards and sanctions in provision of public goods. Economic Inquiry 45, 671–690.

Sutter, M., Haigner, S., Kocher, M.G., 2010. Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. Review of Economics Studies 77 (4), 1540–1566.

Sutter, M., Weck-Hannemann, H., 2004. An experimental test of the public-goods crowding-out hypothesis when taxation is endogenous. Finanzarchiv 60 (1), 94–110.

Tyran, J.R., Feld, L.P., 2006. Achieving compliance when legal sanctions are non-deterrent. The Scandinavian Journal of Economics 108 (1), 135–156.

Warr, P.G., 1982. Pareto optimal redistribution and private charity. Journal of Public Economics 19 (1), 131–138.

Ziegelmeyer, A., Schmelz, K., Ploner, M., 2012. Hidden costs of control: four repetitions and an extension. Experimental Economics 15 (2), 323–340.

Appendix A: Instructions for the exogenous treatment

Welcome to the experiment and thank you for participating!

Please do not talk to other participants.

You are now taking part in an economics experiment. If you read the following instructions carefully, you can – depending on your decisions – earn money in addition to the 50 kronor that you will receive in any case. The entire amount of money that you earn with your decisions will be added up and paid to you in cash at the end of the experiment. These instructions are solely for your private information. **You are not allowed to communicate during the experiment.** Violation of this rule will lead to exclusion from the experiment and all payments. If you have questions, please ask us. We are happy to answer your questions in private.

We will not speak of Swedish kronor during the experiment, but rather of points. Your whole income will first be calculated in points. At the end of the experiment, the total amount of points you earned will be converted to kronor at the following rate:

1 point = 2 kronor

All participants will be divided into groups of three. **Except for us – the experimenters – no one will know who is in which group.** This means that the other participants will neither during nor after the experiment learn how much you earn. We never link names and data from experiments. Please return the form with your ID number in exchange for your payment on your way out after the experiment.

We describe the exact experiment process below. The experiment consists of three parts. You will receive instructions for a part after the previous part has ended. The parts of the experiment are completely independent; the decisions made in one part will not affect your earnings in later parts. The sum of earnings from the different parts will constitute your total earnings from the experiment (together with the 50 kronor show-up fee).

Throughout the experiment, you will write down your decisions on separate pieces of paper (the decision forms), **which will be handed out during the experiment**. The forms will be collected by an experimenter once all participants have made their decisions.

Part I

<u>The decision situation</u>
You will learn how the experiment will be conducted later. Let us first introduce you to the basic decision situation. At the end of the description of the decision situation, you will find control questions that will help you understand the decision situation. You will be a member of a group consisting of **3 people**. Each group member has to decide on the allocation of 20 points. You can put the 20 points into your **private account** or you can put them **fully or partially** into a **group account**. You have to use your entire endowment (= 20 points), which means that the points you put into the group account

and the points you put into the private account have to sum to 20.

Your income from the private account:
You will earn one point for each point you put into your private account. For example, if you put 20 points into your private account (and therefore do not put anything into the group account), your income will amount to exactly 20 points out of your private account. If you put 6 points into your private account, your income from this account will be 6 points. No one except you earns something from your private account.

Your income from the group account:
Each group member will profit equally from the amount you put into the group account. Moreover, you will also get a payoff from the other group members' payments into the group account. The income for each group member out of the group account will be determined as follows:

> *Income from group account =*
> *Sum of all group members' contributions to the group account × 0.5*

If, for example, the sum of all group members' contributions to the group account is 60 points, then you and the other members of your group will each earn 60×0.5 = 30 points out of this account. If the three group members contribute a total of 10 points to the group account, you and the other members of your group will each earn 10×0.5 = 5 points out of this account.

Total income:
Your total income is the sum of your income from your private account and that from the group account:

> *Income from your private account (= 20 – contribution to group account)*
> *+ Income from group account (= 0.5 × sum of all contributions to group account)*
> *= Total income*

Control questions

Please answer the following control questions. They will help you gain an understanding of the calculation of your income, which will vary with how you choose to distribute your 20 points. *Please answer all the questions and write down your calculations.*

1. Each group member has 20 points. Assume that no one in the group (including you) contributes anything to the group account.
What will *your* total income be?   _____
What will the total income of the *other* group members be?   _____

2. Each group member has 20 points. You contribute 20 points to the group account. Each of the other two members of the group also contributes 20 points to the group

account.
What will *your* total income be?   _____
What will the total income of the *other* group members be?   _____

3. Each group member has 20 points. The other 2 members contribute a total of 30 points to the group account.
    a) What will *your* total income be, if you – in addition to the 30 points – contribute 0 points to the group account?
       ***Your* Income**   _____
    b) What will *your* total income be, if you – in addition to the 30 points – contribute 8 points to the group account?
       ***Your* Income**   _____
    c) What will *your* total income be, if you – in addition to the 30 points – contribute 15 points to the group account?
       ***Your* Income**   _____

4. Each group member has 20 points. Assume that you contribute 8 points to the group account.
    a) What is *your* total income if the other group members – in addition to your 8 points – contribute another 7 points to the group account?
       ***Your* Income**   _____
    b) What is *your* total income if the other group members – in addition to your 8 points – contribute another 12 points to the group account?
       ***Your* Income**   _____
    c) What is *your* income if the other group members – in addition to your 8 points – contribute another 22 points to the group account?
       ***Your* Income**   _____

Procedure of Part I [handed out after the completion of the control questions]

Part I includes the decision situation just described to you. The decisions in Part I will only be made **once**. As you know, you will have 20 points at your disposal. You can put them into your private account or you can put them into the group account. Each group member has to make **two types** of contribution decisions, which we will refer to below as the **unconditional contribution** and the **contribution table**.

- In the **unconditional contribution** case you will decide how many of the 20 points you want to put into the group account. Please write down your unconditional contribution in the box on the relevant decision form. Please insert integer numbers only. Your contribution to the private account is determined automatically by the difference between 20 and your contribution to the group account. The decision form looks as follows:

Your contribution to the group account is:	

- On the next decision form you will be asked to complete a **contribution table**. In the contribution table, please indicate **how much <u>you</u> would like to contribute to the group account for each possible average contribution of the <u>other</u> group members** (rounded to the next integer). Thus, you can condition your contribution

on the other group members' average contribution. The contribution table looks as follows:

(Rounded) Average contribution of the other group members to the group account	Your contribution to the group account is
0	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

The numbers in the left column are the possible (rounded) average contributions of the other group members

The numbers in the left column are the possible (rounded) average contributions of the **other** group members to the group account. You will simply have to insert into each input box how many points to contribute to the group account – conditional on the indicated average contribution. **You have to insert a value in each input box**. For example, you will have to indicate how much to contribute to the group account if the others contribute 0 points to the group account on average, how much to contribute if the others contribute 1, 2, or 3 points on average, etc. You can insert any integer number from 0 to 20 in each input box.

Only **the contribution table** will be the payoff-relevant decision for **one randomly determined subject** in each group. Only the **unconditional contribution** will be the payoff-relevant decision for the **other two group members**. Except for us – the experimenters – no one will know who the randomly selected subject is. You will therefore have to think carefully about both types of decisions since either one may become relevant for you. Two examples should make this clear.

Example 1: Assume that **you are the randomly selected subject** in your group. **This implies that your relevant decision will be your contribution table**. The

unconditional contribution is the relevant decision for the other two group members. Assume they made unconditional contributions of 0 and 4 points. The average rounded contribution of these three group members, therefore, is 2 points ((0+4)/2 = 2). If you indicated in your contribution table that you will contribute 1 point to the group account if the others contribute 2 points on average, then the total contribution to the group account is given by 0+4+1 = 5 points. All group members will therefore earn 0.5×5 = 2.5 points out of the group account plus their respective income from the private account. If, instead, you indicated in your contribution table that you would contribute 19 points if the others contribute two points on average, then the total contribution of the group to the group account is given by 0+4+19 = 23. All group members will therefore earn 0.5×23 = 11.5 points out of the group account plus their respective income from the private account.

Example 2: Assume that **you are not the randomly selected subject** in your group**,** implying that **the unconditional contribution is taken as the payoff-relevant decision** for you and one other group member. Assume that your unconditional contribution to the group account is 16 points and that of the other group member is 20 points. The average unconditional contribution of you and the other group member is therefore 18 points ((16+20)/2 = 18). If the group member who is the randomly selected subject indicates in her contribution table that she will contribute 1 point to the group account if the other three group members contribute on average 18 points, then the total contribution to the group account is given by 16+20+1 = 37 points. All group members will therefore earn 0.5×37 = 18.5 points out of the group account plus their respective income from the private account. If, instead, the randomly selected group member indicates in her contribution table that she would contribute 19 points to the group account if the others contribute on average 18 points, then the total contribution to the group account is given by 16+20+19 = 55 points. All group members will therefore earn 0.5×55 = 27.5 points out of the group account plus their respective income from the private account.

The following figure illustrates the situation in example 1. You are the person on the right side (group member number 3). You are the randomly selected subject and therefore your conditional contribution is payoff-relevant. For the other two group members, the unconditional contribution is payoff-relevant.

You will make all your decisions only **once**. After the end of Part I you will get the instructions for Part II.

[Decision sheets handed out]

Part II [handed out after the completion of part I]

We will now conduct another experiment. As in the previous experiment, you are in anonymous **groups of three people**. The members of your group will, however, be **different** in Part II from your group members in Part I.

The **basic decision situation** is the same as that described on page 2 of the instructions of the previous experiment (Part I). Each group member has to decide on the allocation of 20 points. You can put these 20 points into your private account or you can put them fully or partially into a group account. You have to use your entire endowment (= 20 points). The points you put into the group account and the points you put into the private account have to sum to 20. Your income will be determined in the same way as before. Remember:

> *Income from your private account (= 20 – contribution to group account)*
> *+ Income from group account (= 0.5 × sum of all contributions to group account)*
> *= Total income*

In the experiment, each group member is **required to make a minimum contribution of 5 points to the group account**. This means that you must allocate anything from 5 to 20 points to the group account; the remaining points must be allocated to the private account.

Procedure of Part II

Part II includes the decision situation just described to you. The decisions in Part II will only be made **once**. Each group member is required to make a minimum contribution of 5 points to the group account. As in Part I, each group member has to make **two types** of contribution decisions which we will refer to below as the **unconditional contribution** and the **contribution table**.

In the **unconditional contribution** case you decide how many of the 20 points you want to put into the group account. **A minimum of 5 points is required**. In the **contribution table**, please indicate how much you want to contribute to the group account for each possible average contribution of the other group members (rounded to the next integer). The contribution table looks as follows:

(Rounded) Average contribution of the other group members to the group account	Your contribution to the group account is (**a minimum of 5 points is required**)
5	
6	

7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

The numbers in the left column are the possible (rounded) average
contributions of the other group members

The numbers in the left column are the possible (rounded) average contributions of the
other group members to the group account. Please insert into each input box how many
points you want to contribute to the group account – conditional on the indicated
average contribution. **You have to enter at least 5 points in each input box**. For
example, you will have to indicate how much to contribute to the group account if the
others contribute 5 points to the group account on average, how much to contribute if
the others contribute 6, 7, or 8 points on average, etc. You can insert **any integer
number from 5 to 20 in each input box**.

As in Part I, only **the contribution table** will be the payoff-relevant decision for **one
randomly determined subject** in each group. Only the **unconditional contribution**
will be the payoff-relevant decision for the **other two group members**. Except for us –
the experimenters – no one will know who the randomly selected subject is. You will
therefore have to think carefully about both types of decisions since either one can
become relevant for you.

You will make all your decisions only **once**. After the end of Part II you will get the
instructions for Part III.

[Decision sheets handed out]

Part III [handed out after the completion of part II]

We will now ask you to indicate your **expectation** about the average **unconditional
contribution** to the group account (rounded to the next integer) by the **other two group**
members. You will be paid for the accuracy of your expectation:

- If your expectation is exactly right (that is, if your expectation is **exactly** the
  same as the actual average contribution of the other group members), you will
  get **3 points** in addition to your other income from the experiment.
- If your expectation deviates by 1 point from the correct result, you will get 2
  additional points.

- A deviation of 2 points still gives you 1 additional point.
- If your expectation deviates by 3 or more points from the correct result, you will get no additional points.

Each group member has to make two choices (each one pays a maximum of 3 points):

- The expectation about the average unconditional contribution to the group account by the other two group members in **Part I** (when everyone could decide freely on their contribution).
- The expectation about the average unconditional contribution to the group account by the other two group members in **Part II** (when everyone was required to make a minimum contribution of 5 points).

The whole experiment will be finished once Part III has been completed, but we will ask you to fill in a post-experiment questionnaire.

[Decision sheets handed out]

Appendix B: Instructions for the endogenous treatment

Welcome to the experiment and thank you for participating!

Please do not talk to other participants.

You are now taking part in an economics experiment. If you read the following instructions carefully, you can – depending on your decisions – earn money in addition to the 50 kronor that you will receive in any case. The entire amount of money that you earn with your decisions will be added up and paid to you in cash at the end of the experiment. These instructions are solely for your private information. **You are not allowed to communicate during the experiment.** Violation of this rule will lead to exclusion from the experiment and all payments. If you have questions, please ask us. We are happy to answer your questions in private.

We will not speak of Swedish kronor during the experiment, but rather of points. Your whole income will first be calculated in points. At the end of the experiment, the total amount of points you earned will be converted to kronor at the following rate:

1 point = 2 kronor

All participants will be divided into groups of three. **Except for us – the experimenters – no one will know who is in which group.** This means that the other participants will neither during nor after the experiment learn how much you earn. We never link names and data from experiments. Please return the form with your ID number in exchange for your payment on your way out after the experiment.

We describe the exact experiment process below. The experiment consists of three parts. You will receive instructions for a part after the previous part has ended. The parts of the experiment are completely independent; the decisions made in one part will not affect your earnings in later parts. The sum of earnings from the different parts will constitute your total earnings from the experiment (together with the 50 kronor show-up fee).

Throughout the experiment, you will write down your decisions on separate pieces of paper (the decision forms), **which will be handed out during the experiment**. The forms will be collected by an experimenter once all participants have made their decisions.

Part I

The decision situation
You will learn how the experiment will be conducted later. Let us first introduce you to the basic decision situation. At the end of the description of the decision situation, you will find control questions that will help you understand the decision situation. You will be a member of a group consisting of **3 people**. Each group member has to decide on the allocation of 20 points. You can put the 20 points into your **private account** or you can put them **fully or partially** into a **group account**. You have to use your entire endowment (= 20 points), which means that the points you put into the group account and the points you put into the private account have to sum to 20.

Your income from the private account:

You will earn one point for each point you put into your private account. For example, if you put 20 points into your private account (and therefore do not put anything into the group account), your income will amount to exactly 20 points out of your private account. If you put 6 points into your private account, your income from this account will be 6 points. No one except you earns something from your private account.

Your income from the group account:

Each group member will profit equally from the amount you put into the group account. Moreover, you will also get a payoff from the other group members' payments into the group account. The income for each group member out of the group account will be determined as follows:

Income from group account =
Sum of all group members' contributions to the group account × 0.5

If, for example, the sum of all group members' contributions to the group account is 60 points, then you and the other members of your group will each earn 60×0.5 = 30 points out of this account. If the three group members contribute a total of 10 points to the group account, you and the other members of your group will each earn 10×0.5 = 5 points out of this account.

Total income:

Your total income is the sum of your income from your private account and that from the group account:

Income from your private account (= 20 – contribution to group account)
+ Income from group account (= 0.5 × sum of all contributions to group account)
= Total income

Control questions

Please answer the following control questions. They will help you gain an understanding of the calculation of your income, which will vary with how you choose to distribute your 20 points. *Please answer all the questions and write down your calculations.*

1. Each group member has 20 points. Assume that no one in the group (including you) contributes anything to the group account.
What will *your* total income be? _____
What will the total income of the *other* group members be? _____

2. Each group member has 20 points. You contribute 20 points to the group account. Each of the other two members of the group also contributes 20 points to the group account.

What will *your* total income be? _____
What will the total income of the *other* group members be? _____

3. Each group member has 20 points. The other 2 members contribute a total of 30 points to the group account.
   a) What will *your* total income be, if you – in addition to the 30 points – contribute 0 points to the group account?
      ***Your* Income** _____
   b) What will *your* total income be, if you – in addition to the 30 points – contribute 8 points to the group account?
      ***Your* Income** _____
   c) What will *your* total income be, if you – in addition to the 30 points – contribute 15 points to the group account?
      ***Your* Income** _____

4. Each group member has 20 points. Assume that you contribute 8 points to the group account.
   a) What is *your* total income if the other group members – in addition to your 8 points – contribute another 7 points to the group account?
      ***Your* Income** _____
   b) What is *your* total income if the other group members – in addition to your 8 points – contribute another 12 points to the group account?
      ***Your* Income** _____
   c) What is *your* income if the other group members – in addition to your 8 points – contribute another 22 points to the group account?
      ***Your* Income** _____

Procedure of Part I [handed out after the completion of the control questions]

Part I includes the decision situation just described to you. The decisions in Part I will only be made **once**. As you know, you will have 20 points at your disposal. You can put them into your private account or you can put them into the group account. Each group member has to make **two types** of contribution decisions, which we will refer to below as the **unconditional contribution** and the **contribution table**.

- In the **unconditional contribution** case you will decide how many of the 20 points you want to put into the group account. Please write down your unconditional contribution in the box on the relevant decision form. Please insert integer numbers only. Your contribution to the private account is determined automatically by the difference between 20 and your contribution to the group account. The decision form looks as follows:

Your contribution to the group account is:	

- On the next decision form you will be asked to complete a **contribution table**. In the contribution table, please indicate **how much you would like to contribute to the group account for each possible average contribution of the other group members** (rounded to the next integer). Thus, you can condition your contribution on the other group members' average contribution. The contribution table looks as

32

follows:

(Rounded) Average contribution of the other group members to the group account	Your contribution to the group account is
0	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

The numbers in the left column are the possible (rounded) average
contributions of the other group members

The numbers in the left column are the possible (rounded) average contributions of the **other** group members to the group account. You will simply have to insert into each input box how many points to contribute to the group account – conditional on the indicated average contribution. **You have to insert a value in each input box**. For example, you will have to indicate how much to contribute to the group account if the others contribute 0 points to the group account on average, how much to contribute if the others contribute 1, 2, or 3 points on average, etc. You can insert any integer number from 0 to 20 in each input box.

Only **the contribution table** will be the payoff-relevant decision for **one randomly determined subject** in each group. Only the **unconditional contribution** will be the payoff-relevant decision for the **other two group members**. Except for us – the experimenters – no one will know who the randomly selected subject is. You will therefore have to think carefully about both types of decisions since either one may become relevant for you. Two examples should make this clear.

Example 1: Assume that **you are the randomly selected subject** in your group. **This implies that your relevant decision will be your contribution table**. The unconditional contribution is the relevant decision for the other two group members.

Assume they made unconditional contributions of 0 and 4 points. The average rounded contribution of these three group members, therefore, is 2 points ((0+4)/2 = 2). If you indicated in your contribution table that you will contribute 1 point to the group account if the others contribute 2 points on average, then the total contribution to the group account is given by 0+4+1 = 5 points. All group members will therefore earn 0.5×5 = 2.5 points out of the group account plus their respective income from the private account. If, instead, you indicated in your contribution table that you would contribute 19 points if the others contribute two points on average, then the total contribution of the group to the group account is given by 0+4+19 = 23. All group members will therefore earn 0.5×23 = 11.5 points out of the group account plus their respective income from the private account.

Example 2: Assume that **you are not the randomly selected subject** in your group**,** implying that **the unconditional contribution is taken as the payoff-relevant decision** for you and one other group member. Assume that your unconditional contribution to the group account is 16 points and that of the other group member is 20 points. The average unconditional contribution of you and the other group member is therefore 18 points ((16+20)/2 = 18). If the group member who is the randomly selected subject indicates in her contribution table that she will contribute 1 point to the group account if the other three group members contribute on average 18 points, then the total contribution to the group account is given by 16+20+1 = 37 points. All group members will therefore earn 0.5×37 = 18.5 points out of the group account plus their respective income from the private account. If, instead, the randomly selected group member indicates in her contribution table that she would contribute 19 points to the group account if the others contribute on average 18 points, then the total contribution to the group account is given by 16+20+19 = 55 points. All group members will therefore earn 0.5×55 = 27.5 points out of the group account plus their respective income from the private account.

The following figure illustrates the situation in example 1. You are the person on the right side (group member number 3). You are the randomly selected subject and therefore your conditional contribution is payoff-relevant. For the other two group members, the unconditional contribution is payoff-relevant.

You will make all your decisions only **once**. After the end of Part I you will get the instructions for Part II.

[Decision sheets handed out]

Part II [handed out after the completion of part I]

We will now conduct another experiment. As in the previous experiment, you are in anonymous **groups of three people**. The members of your group will, however, be **different** in Part II from your group members in Part I.

The **basic decision situation** is the same as that described on page 2 of the instructions of the previous experiment (Part I). Each group member has to decide on the allocation of 20 points. You can put these 20 points into your private account or you can put them fully or partially into a group account. You have to use your entire endowment (= 20 points). The points you put into the group account and the points you put into the private account have to sum to 20. Your income will be determined in the same way as before. Remember:

Income from your private account (= 20 – contribution to group account)
+ Income from group account (= 0.5 × sum of all contributions to group account)
= Total income

In the experiment, you and the other group members can participate in a vote and make other decisions. There are **two possible institutions**.

Institution (a) – Decide freely on your contribution
Under this institution, each group member is free to decide how to allocate their 20 points. This means that you can allocate anything from 0 to 20 points to the group account; the remaining points must be allocated to the private account. This is the same decision situation as in Part 1.

Institution (b) – Required to make a minimum contribution of 5 points
Under this institution, each group member is required to make a minimum contribution of 5 points to the group account. This means that you must allocate anything from 5 to 20 points to the group account; the remaining points must be allocated to the private account.

A vote
You and all other group members will participate in a vote on the institution that will be valid in your group. You have to indicate **which one** of the two possible institutions you prefer for your group. The institution that receives a majority of the votes (i.e., 2 or 3 votes) will be valid in your group.

Procedure of Part II

Part II includes the decision situation just described to you. The decisions in Part II will only be made **once.** First each group member will participate in the vote. After you have

learned the outcome of the vote, please make your contribution decision under the institution that is valid in your group. You are asked to make both an unconditional and a conditional contribution, but **for this institution only**.

1. The vote
Please vote by **marking the box next to the institution you prefer** for your group. The voting form looks as follows:

I vote for Institution (a) – everyone can decide freely on their contribution:	☐
I vote for Institution (b) – everyone is required to make a minimum contribution of 5 points:	☐

2. Your contribution decision
You will **learn the outcome of the vote** in your group when the decision forms for the contribution are handed out.

- **If a majority of the members of your group have voted for Institution (a)**, you and everyone else in your group will receive decision forms for **this institution only**. Each group member is free to decide how to allocate their 20 points. The forms look exactly as in Part I. In the unconditional contribution case, you must decide how many of the 20 points you want to put into the group account. In the contribution table, please indicate how much you want to contribute to the group account for each possible average contribution of the other group members (rounded to the next integer).

- **If a majority of the members of your group have voted for Institution (b)**, you and everyone else in your group will receive decision forms for **this institution only**. Each group member is required to make a minimum contribution of 5 points to the group account. In the unconditional contribution case, you must decide how many of the 20 points you want to put into the group account. **A minimum of 5 points is required**. In the contribution table, please indicate how much you want to contribute to the group account for each possible average contribution of the other group members (rounded to the next integer). The contribution table looks as follows:

(Rounded) Average contribution of the other group members to the group account	Your contribution to the group account is (**a minimum of 5 points is required**)
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	

15	
16	
17	
18	
19	
20	

The numbers in the left column are the possible (rounded) average
contributions of the other group members

The numbers in the left column are the possible (rounded) average contributions of the
other group members to the group account. Please insert into each input box how many
points you want to contribute to the group account – conditional on the indicated
average contribution. **You have to enter at least 5 points in each input box**. For
example, you will have to indicate how much to contribute to the group account if the
others contribute 5 points to the group account on average, how much to contribute if
the others contribute 6, 7, or 8 points on average, etc. You can insert **any integer
number from 5 to 20 in each input box**.

As in Part I, only **the contribution table** will be the payoff-relevant decision for **one
randomly determined subject** in each group. Only the **unconditional contribution**
will be the payoff-relevant decision for the **other two group members**. Except for us –
the experimenters – no one will know who the randomly selected subject is. You will
therefore have to think carefully about both types of decisions since either one can
become relevant for you.

You will make all your decisions only **once**. After the end of Part II you will get the
instructions for Part III.

[Decision sheets handed out; first for the vote then for the contribution decisions for the
institution implemented by the vote in each group]

Part III [handed out after the completion of part II]

We will now ask you to indicate your **expectation** about the **average unconditional
contribution** to the group account (rounded to the next integer) by the **other two group**
members. You will be paid for the accuracy of your expectation:

- If your expectation is exactly right (that is, if your expectation is **exactly** the
  same as the actual average contribution of the other group members), you will
  get **3 points** in addition to your other income from the experiment.
- If your expectation deviates by 1 point from the correct result, you will get 2
  additional points.
- A deviation of 2 points still gives you 1 additional point.
- If your expectation deviates by 3 or more points from the correct result, you will
  get no additional points.

Each group member has to make two choices (each one pays a maximum of 3 points):

- The expectation about the average unconditional contribution to the group
  account by the other two group members in **Part I** (when everyone could decide

freely on their contribution).

- The expectation about the average unconditional contribution to the group account by the other two group members in **Part II** (when the outcome of the vote determined whether everyone should be allowed to decide freely on their contribution, or if everyone should be required to make a minimum contribution of 5 points).

The whole experiment will be finished once Part III has been completed, but we will ask you to fill in a post-experiment questionnaire.

[Decision sheets handed out]

Tables

Table 1. Results from the public goods experiments: unconditional contributions.

Treatment	Standard	MCL[a]	H_0: Standard=MCL (p-values)
Exogenous ($n = 51$)	8.35 (6.80)	9.43 (5.47)	0.01
Endogenous ($n = 93$)	8.24 (7.00)	9.09 (5.68)	< 0.01

Note: Standard deviations in brackets. n = number of subjects. [a] For subjects in groups rejecting the endogenous minimum level, the minimum level is not imposed in the second stage; instead they participate in a standard public goods experiment.

Table 2. Voting behavior in the endogenous treatment.

	Standard contr.	Standard belief	MCL contr.	MCL belief	H_0: Standard contr.=MCL contr. (p-values)	H_0: Standard contr.= Standard belief (p-values)
Exogenous ($n = 51$)	8.35	7.76	9.43	9.57	0.01	< 0.01
Endogenous ($n = 93$)	8.25	7.66	9.09	9.01	< 0.01	< 0.01
Endogenous: own and group vote						
-Yes/Yes ($n = 72$)	8.39	7.43	9.75	9.54	< 0.01	< 0.01
- No/Yes ($n = 9$)	8.22	8.89	6.67	7.78	0.63	0.37
- No/No ($n = 9$)	8.44	9.22	7.78	7.78	0.16	0.58
- Yes/No ($n = 3$)	4.33	4.67	4.33	3.67	.	0.17

Table 3. Behavioral response to imposing the minimum level.

	Higher in MCL	No difference	Lower in MCL
Exogenous ($n = 51$)	49%	31%	20%
Endogenous ($n = 81$)	43%	40%	17%
Endogenous: own and group vote			
-Yes/Yes ($n = 72$)	44%	42%	14%
- No/Yes ($n = 9$)	33%	22%	44%

Table 4. Behavioral response to imposing the minimum level for exogenous treatment ($n = 51$).

Contributions in standard public goods experiment	Contributions in MCL public goods experiment	Proportion of subjects	Average contribution in standard public goods experiment	Average contribution in minimum level public goods experiment
0,…,4	5	23%	0.00	5.00
0,…,4	6,…,20	2%	2.00	6.00
5	5	6%	5.00	5.00
5	6,…,20	8%	5.00	8.25
6,…,20	5	12%	8.83	5.00
6,…,20	6,…,20	49%	13.44	13.48

Table 5. Behavioral response to imposing the minimum level for endogenous treatment (yes-groups, $n = 81$).

Contributions in standard public goods experiment	Contributions in MCL public goods experiment	Proportion of subjects	Average contribution in standard public goods experiment	Average contribution in minimum level public goods experiment
0,…,4	5	27%	0.10	5.00
0,…,4	6,…,20	2%	1.50	7.50
5	5	5%	5.00	5.00
5	6,…,20	5%	5.00	8.25
6,…,20	5	10%	11.38	5.00
6,…,20	6,…,20	51%	13.22	13.27

Table 6. Cooperative preferences by treatment (linear regression).

Dependent variable: Conditional contribution	(1)	(2)
Constant	1.37 (0.40)***	1.53 (0.75)**
Others	0.45 (0.04)***	0.44 (0.06)***
MCL	2.57 (0.34)***	2.05 (0.64)***
MCL × Others	-0.07 (0.03)**	-0.03 (0.04)
Endo	-	-0.26 (0.88)
Endo × Others	-	-0.02 (0.08)
Endo × MCL	-	0.84 (0.74)
Endo × MCL × Others	-	-0.06 (0.06)
Observations	4889	4889
- Clusters	132	132

Note: Robust standard errors clustered on individuals (in brackets). *Others* is the average contribution of the others in the group, *MCL* is a dummy for the second stage when the minimum level is imposed, and *Endo* is a dummy for the endogenous treatment. The dependent variable is individual i's conditional contribution, both in the baseline and with the minimum provision level. *** denotes significance at the 1% level, ** significance at the 5% level, and * significance at the 10% level.

Table 7. Distribution of contributor types for exogenous treatment ($n = 51$).

Standard \ MCL	Conditional cooperators	Free riders (=5)	Hump-shaped	Others	
Conditional cooperators	50.98%	5.88%	0.00%	7.84%	64.70%
Free riders (=0) (=5)	0.00%	17.65%	0.00%	0.00%	17.65%
Hump-shaped	0.00%	0.00%	0.00%	1.96%	1.96%
Others	1.96%	1.96%	0.00%	11.76%	15.69%
	52.94%	25.49%	0.00%	21.57%	100.00%

Note: Type definition follows Fischbacher et al. (2001). Conditional cooperators: increasing schedule or positive Spearman correlation (with others' average contribution) at 1%; free riders: contribute the lowest amount allowed; hump-shaped: positive (negative) Spearman correlation at 1% up to (beyond) their highest contribution.

Table 8. Distribution of contributor types for endogenous treatment (yes-groups, $n = 81$).

Standard \ MCL	Conditional cooperators	Free riders (=5)	Hump-shaped	Others	
Conditional cooperators	41.98%	9.88%	1.23%	6.17%	59.26%
Free riders (=0) (=5)	0.00%	13.58%	0.00%	1.23%	14.81%
Hump-shaped	2.47%	1.23%	2.47%	1.23%	7.41%
Others	2.47%	6.17%	0.00%	9.88%	18.52%
	46.91%	30.86%	3.70%	18.51%	100.00%

Note: Type definition follows Fischbacher et al. (2001). Conditional cooperators: increasing schedule or positive Spearman correlation (with others' average contribution) at 1%; free riders: contribute the lowest amount allowed; hump-shaped: positive (negative) Spearman correlation at 1% up to (beyond) their highest contribution.

Table 9. Distribution of contributor types when a minimum level is manually imposed on the data from the standard public goods experiment.

	Exogenous ($n = 51$)		Endogenous (yes-groups, $n = 81$)	
	Standard	MCL	Standard	MCL
Conditional cooperators	56.86%	52.94%	54.32%	46.91%
Free riders (=5)	25.49%	25.49%	27.16%	30.86%
Hump-shaped	1.96%	0.00%	6.17%	3.70%
Others	15.69%	21.57%	12.35%	18.52%

Note: Type definition follows Fischbacher et al. (2001). In this table we define contributor types in the standard public goods game are after we manually adjust all subjects up to the minimum of 5 tokens in the standard public goods experiment. This implies that someone who never contributes above 5 tokens in the standard public goods experiment is characterized as a free rider.

Figures

Figure 1. Individual contributions in exogenous treatment ($n = 51$).

Figure 2. Individual contributions in endogenous treatment ($n = 81$).

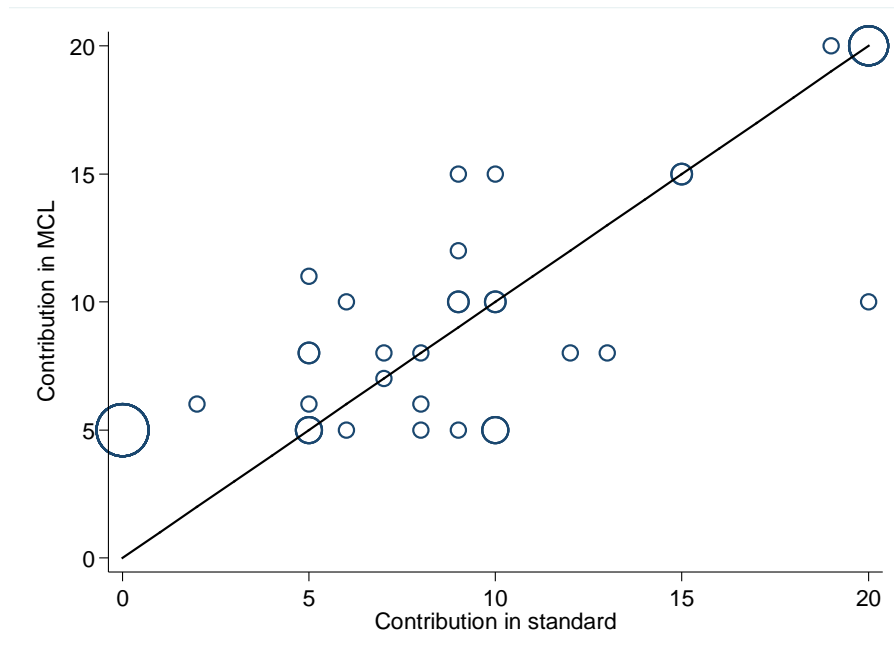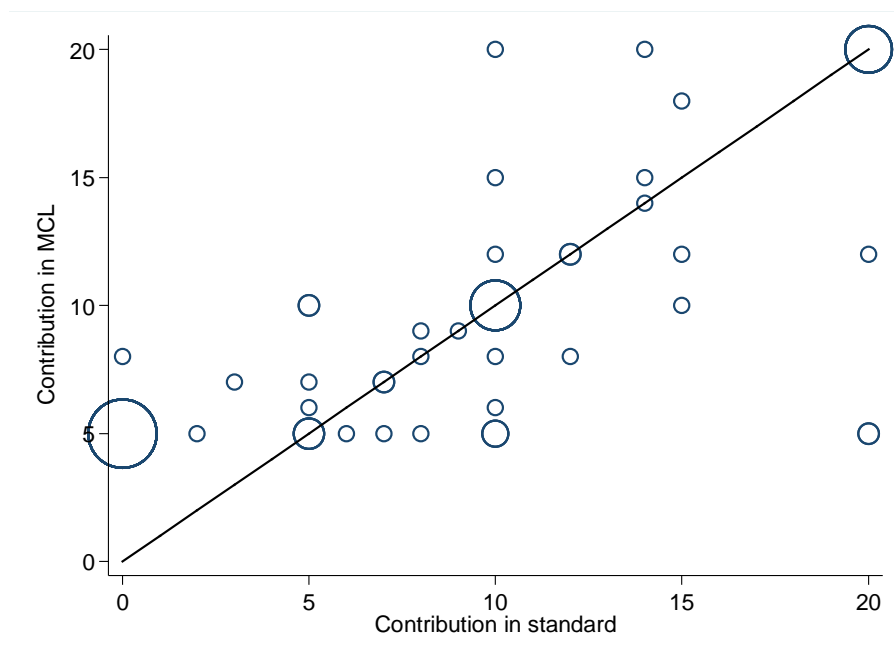**Figure 3.** Contributions and beliefs about others' contributions for exogenous treatment ($n = 51$).
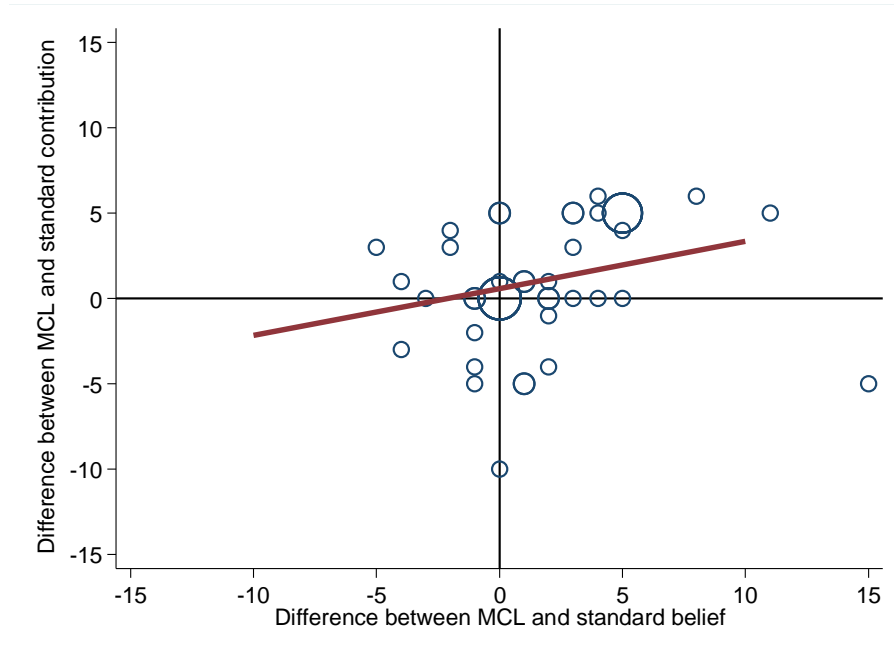

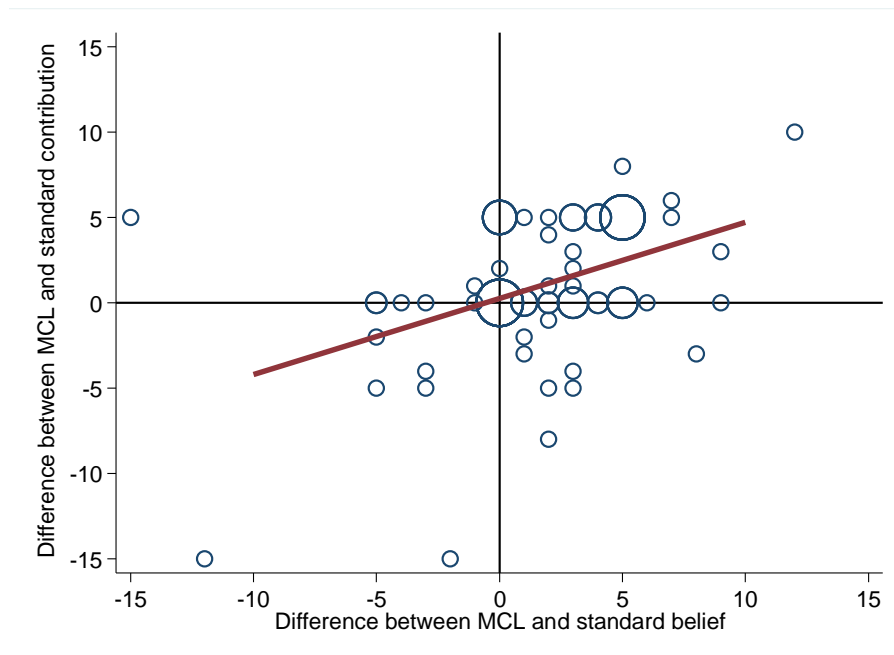**Figure 4.** Contributions and beliefs about others' contributions for endogenous treatment ($n = 81$).

Figure 5. Average conditional contributions for exogenous treatment ($n = 51$).
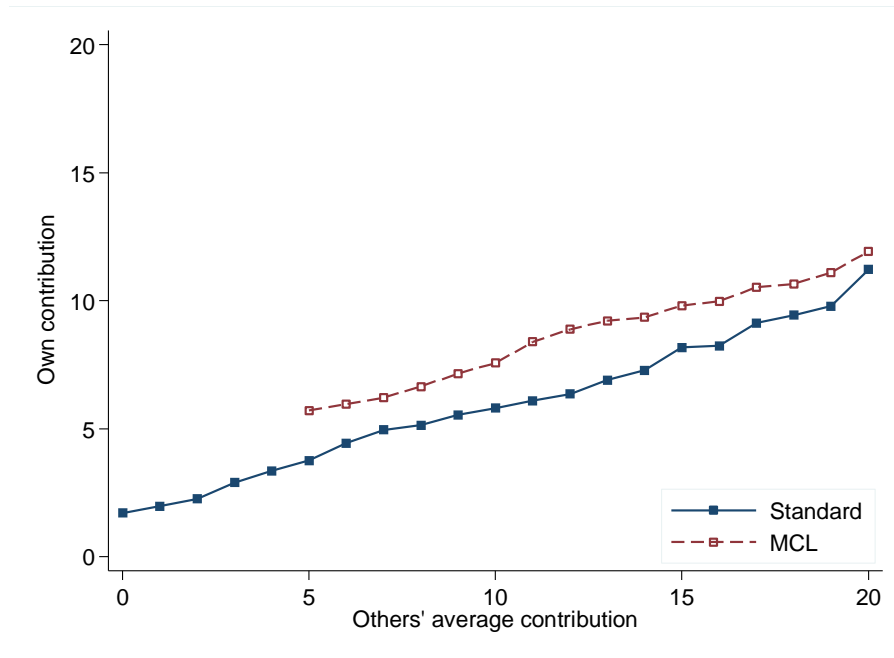
Figure 6. Average conditional contributions for endogenous treatment ($n = 81$).
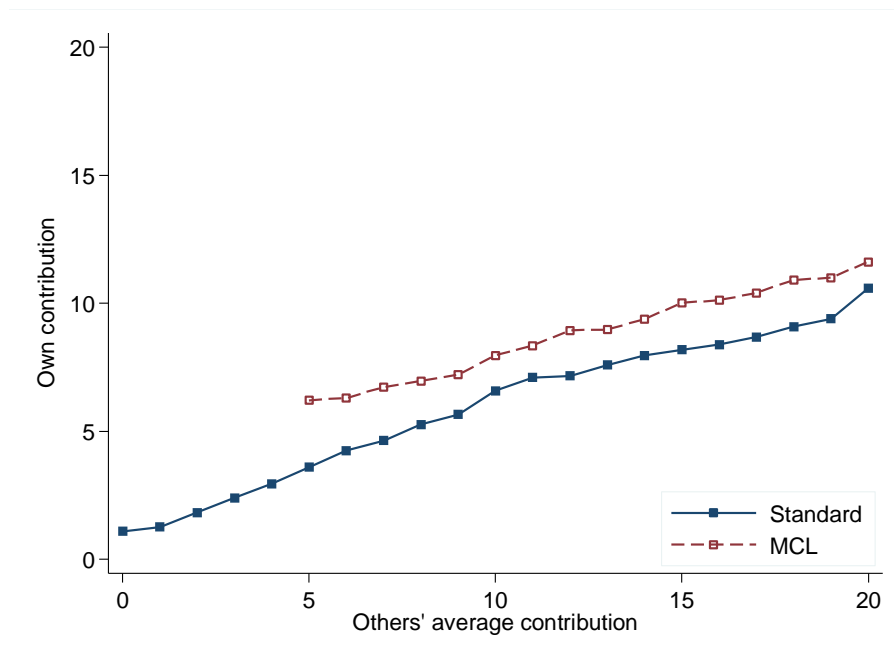
Paper II

Framing and Minimum Levels in Public Good Provision

Peter Martinsson, Haileselassie Medhin, and Emil Persson[*]

Department of Economics, University of Gothenburg, Sweden

Abstract

Using a laboratory experiment in the field, we examine how the choice architecture of framing a social dilemma – *give to* or *take from* a public good – interacts with a policy intervention that enforces a minimum contribution level to the public good. We find that cooperation is significantly higher in the give frame than in the take frame in our standard public goods experiment. When a minimum contribution level is introduced, contributions are significantly higher in the take frame since contributions are crowded out in the give frame but crowded in in the take frame. Our results therefore stress the importance of choosing the frame when making policy recommendations.

Keywords: Choice architecture, Framing, Public goods, Minimum level, Experiment, Ethiopia.
JEL Classification: C91, H41.

---

[*] Corresponding author at: Department of Economics, University of Gothenburg, Box 640, SE 405 30 Gothenburg, Sweden; Tel.: +46 31 786 2927.
 E-mail addresses:
 peter.martinsson@economics.gu.se (P. Martinsson),
 haileselassie.medhin@economics.gu.se (H. Medhin),
 emil.persson@economics.gu.se (E. Persson).

1. Introduction

This paper focuses on the design of institutions that foster cooperation in the provision of local public goods, for example local environmental quality. Using lab-in-the-field experiments, we investigate the role of framing and minimum levels in public good provision in rural Ethiopia. These are two prominent institutional factors that policy makers can use in their role as choice architects.[1] Thaler and Sunstein (2009) coined the expression nudging to describe that the policy makers, i.e., choice architects, have the possibility to affect people's behavior and still respect freedom of choice, i.e., libertarian paternalism. The objective of our paper is to investigate the interaction effects between framing and minimum levels in a non-student and non-Western subject pool. Although previous experiments have shown that other institutions such as monetary punishment and exclusions are promising institutions to increase cooperation (e.g., see overview in Chaudhuri, 2011), we firmly believe that framing and minimum level are the institutions that are possible and easy to affect by policy makers, especially in developing countries. To our knowledge, this is the first study that experimentally explores the interaction between framing and formal institutions such as minimum levels in public good games.

Policy makers can affect cooperative behavior by changing the frame of a given cooperation problem, by choosing whether the activity should be designed in terms of giving to or taking from a public good. We frame public goods in terms of giving or taking and investigate the impact on cooperative behavior. Previous experimental literature on framing effects in social dilemmas has mostly focused on understanding the difference in behavior, and the results show more cooperation in give frames, or no difference,[2] despite the fact that individuals face the same fundamental economic problem in both frames (e.g., Andreoni, 1995; Cox, 2015; Dufwenberg, et al., 2011; Fosgaard et al., 2014; Khadjavi and Lange, 2015). However, our focus is on the relevance of framing for policy design, in particular concerning the interaction between framing and minimum levels in public good provision. Minimum levels are often imposed in order to secure some provision of a public good. For example, in developing countries it is common to provide local public goods directly through mandatory community-based labor exchange, with provision of the public good enforced without a

---

[1] A neutral design does not exist. Any design choice a decision maker makes influence people's behavior, from the order of food choices on a menu to whether CO_2 compensation of a flight is by default included or not included in the flight price.

[2] For a summary of previous experiments on take and give frames, see Gächter et al. (2014). See also Cookson (2000).

formal institution, through peer-pressure and social norms (Olken and Singhal, 2011; Ostrom, 1991). A minimum level should increase the total provision of the public good since any free rider is forced to contribute at least the announced minimum, and this in turn might make conditional cooperators increase their contribution since they want to cooperate if others cooperate and vice versa (e.g., Fischbacher et al., 2001). However, the minimum level could also result in crowding out, i.e., that individuals contribute less to the public good when the minimum level is imposed than when it is not imposed because the minimum level might send a signal of distrust to intrinsically cooperative agents. In this case the minimum level entails a "hidden cost of control" (Falk and Kosfeld, 2006; Ziegelmeyer et al., 2012).[3] Results from previously conducted experiments of minimum levels using the standard give frame, i.e., subjects are asked to give to the public good, are mixed regarding the effect of minimum levels on public good provision. For instance, Andreoni (1993) and Gronberg et al. (2012) find a positive effect of a minimum level on public good provision using concave payoff functions, Eckel et al. (2005) implement a dictator game with a charity as recipient and find that a minimum level crowds out donations to the charity when it is framed as a tax, and Kocher et al. (2016) use a linear public goods game and find instances of both crowding out and crowding in following the introduction of a minimum level.

The decision maker has the important role of being a choice architect. In order to use framing as a policy design tool it is helpful to distinguish between situations where it is possible to meaningfully change the frame of a given cooperation problem and situations where the frame is simply a structural aspect of the cooperation problem. In the former case, changing the framing can be a policy tool in itself and hence an important design problem for a choice architect. For instance, cooperative behavior can be encouraged by using positive language (e.g., emphasize the positive effect of doing something which is good for the community) or negative language (e.g., emphasize the negative effect of doing something which is bad for the community), and the policy maker can also affect the design of the activity itself such that it involves either giving to or taking from a public good.[4] In the latter case, understanding framing as a structural aspect of the cooperation problem could be useful in the selection of other policy tools such as minimum levels. It is very important for policy makers to know whether

---

[3] See for instance Deci (1971, 1975) and Bowles and Polanía-Reyes (2012) for general discussions on crowding out of intrinsic motivation and control aversion.

[4] This is closely related to the literature on choice architecture. For instance, framing has been found to affect savings behavior (Brown et al., 2008), energy conservation (Gromet et al., 2013), and the impact of conditional incentives on worker productivity (Hossain and List, 2012) and teacher performance (Fryer et al., 2012).

minimum levels work differently for give and take frames when considering future choice architecture designs.[5] This is indeed the key motivation for exploring the interaction between framing and minimum levels.

Framing can influence how subjects conceive their own actions and others' actions (Tversky and Kahneman, 1981). In terms of our public goods game, actions are kind in the give frame ("to give") but unkind in the take frame ("to take"), and this possibly affects behavior. For instance, we might expect higher public good provision in the give frame because of reciprocity (since subjects reciprocate kind actions with kindness; Cox et al., 2008, 2013); or because of "warm glow," if subjects derive more utility from doing something good than from refraining from doing something bad (Andreoni, 1995).[6] However, if subjects care about what is socially appropriate, then we might instead expect higher provision in the take frame (since it is a less socially appropriate action to take than to give; Krupka and Weber, 2013). The minimum level could have an effect along some or all of these dimensions. It could for instance act as a signal or reference point, thus affecting subjects' view of what is socially appropriate and also their belief about what others are doing. These effects might differ depending on whether public good provision is framed as give or take. Moreover, the minimum level might be perceived as more lenient in the take frame since it is binding from above ("take at most") and thus someone who considers taking less than the imposed maximum might feel unaffected by this constraint; conversely, by contributing more than the imposed minimum in the give frame, one might feel that the part of the contribution stipulated by the mandatory minimum ("give at least") was a forced choice. Thus, the predicted net effects are difficult to pin down theoretically. Our experiments provide empirical evidence on the relationship between framing and minimum levels.

Our results are based on lab-in-the-field experiments in rural Ethiopia, where the subjects were farm household heads actively engaged in a number of public-good-like decisions such as environmental rehabilitation and the maintenance of local infrastructure. Thus, we also contribute to extending the conventional analysis of cooperation to a non-student and non-Western subject pool (e.g., Henrich et al., 2010). This is especially important since societies across the world differ substantially in terms

---

[5] It is possible that both dimensions of framing as policy are applicable to the same cooperation problem. For example, the problem of labor contribution to a local public good can be considered as a structural give frame, while the problem of over-harvesting from a local public good can be considered as a structural take frame. It is however possible to highlight the positive and negative consequences of individuals' actions in each of the above problems.

[6] See also Cubitt et al. (2011), who find that subjects condemn free riding more strongly in give situations than in take situations.

of social organization, trust, fairness norms, and also in the nature of day-to-day cooperation problems.[7] Results from developing-country subject pools like ours could therefore be particularly important both in examining the generalizability of experimental findings and in drawing conclusions about social dilemmas prevalent in such places. Moreover, our results provide information for the choice architecture of public goods projects. Fosgaard et al. (2014) were the first to implement a public goods game framed as either take-from or give-to in a non-student sample, namely a large-scale sample of the Danish population. Interestingly, their treatment effects differ from the majority of experiments using conventional student samples in that they find more cooperation in the take frame. We extend the analysis further by conducting experiments with more than 300 farmers in Ethiopia. We find strong effects of framing and minimum levels on cooperation. More precisely, we find that cooperation is significantly higher in the give frame than in the take frame in our standard public goods experiment. When a minimum contribution level is introduced, contributions are significantly higher in the take frame since contributions are crowded out in the give frame but crowded in in the take frame. Overall, the highest level of cooperation is observed in the give frame with no minimum level. Our results show the importance of framing for choice architects. Furthermore, our results could be important inasmuch as they point to situations where policy makers can increase the efficiency of interventions with simple and cost-efficient framing techniques combined with formal institutions such as minimum levels in public good provision. These interventions are easier to implement than other institutional features such as monetary punishment or exclusion. The rest of the paper is organized as follows. Section 3 describes the experimental design, Section 4 presents the results, and we provide concluding remarks in Section 5.

2. Experiment

2.1. Design

The primary focus of our paper is to investigate the effect of (i) framing phrased as either give to (GIVE) or take from (TAKE) the public good combined with (ii) a minimum level. We begin with a description of the framing of the public goods and then we explain how the minimum level was implemented and finally how we tested for the effect on contributions to the public goods in both frames.

---

[7] See for instance Henrich et al. (2001) for cross-cultural comparisons of economic experiments in small-scale societies. See also Herrmann et al. (2008) and Vieder et al. (2015).

5

We implement a linear public goods experiment based on the design developed by Fischbacher et al. (2001), which in addition to eliciting unconditional contributions to public goods also uses the strategy method to elicit contributions to public goods conditional on others' average contributions.[8,9] We begin with a description of the public goods using the give frame, i.e., how much a subject would like to give[10] to a public good. This is the common way to phrase contributions in public goods experiments. Subjects are matched into groups of four, each with an endowment of 10 Ethiopian birr and the possibility to contribute any integer amount c from 0 to 10 Ethiopian birr to the public good. To facilitate understanding, we choose to use the real currency directly rather than an experimental currency.[11] The marginal per capita return from the public good is 0.5, i.e., a contribution of one unit results in 0.5 units of income for each of the four group members. This can easily be explained to the subjects by explaining that the total amount contributed to the public good is doubled and then split equally between the subjects. Since the marginal per capita return is 0.5 and the social marginal per capita return is 2, it is a dominant strategy for a payoff-maximizing individual to contribute nothing to the public good. However, it is socially optimal to contribute the whole endowment. The payoff for subject i is given by

$$\pi_i = 10 - c_i + 0.5 \sum_{j=1}^{4} c_j. \tag{2}$$

The Fischbacher et al. design employs the strategy method. Subjects make two types of giving decisions, one unconditional and one conditional. In the unconditional decision, subjects decide how much they wish to give to the public good without knowing anything about anybody else's contributions. In the conditional decision, each subject decides on the amount to give conditional on the average amount given by the other three group members. This is implemented by each subject stating how much she/he

---

[8] For a discussion on the validity of using the strategy method to elicit cooperative preferences, see, e.g., Fischbacher et al. (2012). For a general discussion on the policy implications of results from public goods experiments, see Gächter (2007). For other experiments based on this design, see, e.g., Fischbacher and Gächter (2010), Herrmann and Thöni (2009), Kocher et al. (2008), Martinsson et al. (2013), and Martinsson et al. (2015). See also Rustagi et al. (2010) and Kosfeld and Rustagi (2015) for interesting applications to the management of forest commons in Ethiopia.

[9] We chose a design with linear rather than concave payoff functions (e.g., Andreoni, 1993) since it facilitates understanding and the use of the strategy method, which we think is important. This means that the minimum level is set above rather than below the Nash equilibrium of zero contribution, which makes the welfare effect of the minimum level more salient.

[10] To facilitate understanding, we deliberately use the word "give" instead of "contribute" when we write about the GIVE and TAKE treatments.

[11] In our Ethiopian case, this worked particularly well since the value of Ethiopian birr matched well with the opportunity cost we intended to use. The experiment with two stages was calibrated to give on average almost a daily salary (30 birr).

would give to the public good for each possible average amount given (in integers) by the others in the group ranging from 0 to 10.

To make the decisions incentive compatible, the unconditional decision will be payoff relevant for three randomly selected group members, and by using the average unconditional amount given (rounded to the next integer) by them, the amount given by the fourth member is determined as his or her conditional giving matching that specific average amount. Payoffs are then calculated based on these amounts.

In the take frame, the decision is framed as how much a subject would like to take from the public good. We follow standard procedures in the framing literature when we implement this frame (e.g., Andreoni, 1995; Fosgaard et al., 2014). In this treatment, subjects are not endowed with 10 Ethiopian birr as in the give frame. Instead, the public good consists of $4 \times 10 = 40$ Ethiopian birr and subjects decide on the (integer) amount w from 0 to 10 birr that they wish to take from the public good. The payoff function for subject i is given by

$$\pi_i = w_i + 0.5\left(40 - \sum_{j=1}^{4} w_j\right). \tag{3}$$

The incentives are exactly the same across GIVE and TAKE since equations (2) and (3) describe the same underlying payoff function and only differ in how it is framed.

In addition to testing for framing effects, we investigate the effect of introducing a minimum level in each frame. In the give frame, subjects must give at least the announced minimum level of 2 Ethiopian birr, which corresponds to 20% of the endowment. We follow Andreoni (1993) and present the minimum level as a restriction on individuals' choice set. For instance, conditional giving is elicited by subjects making eleven giving decisions (one for each possible integer average amount given by the others in the group); when the minimum level of 2 Ethiopian birr has been implemented, subjects make nine giving decisions instead of eleven and they know that they are not allowed to give less than 2 Ethiopian birr. In the take frame, where the procedures are exactly the same with the exception that the game is presented as take-from rather than give-to the public good, subjects must not take more than an announced amount of 8 Ethiopian birr, and hence leave at least 2 Ethiopian birr in the public good.

In the experiment, subjects participate in sessions with public goods experiments framed as TAKE or GIVE and within each session they complete two stages. The first stage (Baseline) is a standard public goods experiment. It is followed by a second stage (MCL) which is a public goods experiment with an imposed minimum contribution

level.[12] We use stranger matching, i.e., they are re-matched with three new group members in the second stage, and as described above a subject makes two types of decisions (unconditional and conditional) in each stage. They are paid for their decisions in both stages but learn nothing about the decisions of others in the group until they are paid some days after the experiment.

2.2. Procedural details

The experiment was conducted in 2013 in rural Ethiopia. It was a separate module of a household survey on community forestry that covered 15 villages scattered across the four major regions of the country. We ran experiments in eight of these villages.[13] In each region we first randomly selected two villages from each region. Then one of the two villages was randomized into the give treatment and the other into the take treatment.[14] With randomization of treatments at the regional level, we obtain a balanced sample where the regional mix of subjects is similar across the treatments. This is to ensure that our results are not driven by regional differences, which could be quite large given the wide (and diverse) geographical area covered.

In each village, households were randomly selected for participation in the survey. The household heads were interviewed in their respective houses in private by a trained enumerator. At the end of the household survey, they were asked if they would like to participate in an economic experiment. All household heads covered in the survey agreed to participate in the experiment. The experiment was conducted face-to-face similar to in for example Henrich et al. (2001). The above described public goods experiment was clearly described to the subjects, who after a number of comprehension questions completed both stages of the experiment. Subjects were informed that similar experiments had or would take place in other households in their village and that they would be randomly and anonymously matched against other household heads in their village. They were also informed that payments would be distributed on a specific date and that this procedure was the same for everybody else participating in the

---

[12] We do not test for order effects and the reason is that the order we impose is natural in most cases, in the sense that we are interested in the effect of a minimum level when subjects are experienced with the baseline, which is a voluntary contribution mechanism in the form of either take-from or give-to the public good. This is what typically happens in reality: a public goods problem is identified and a remedy is sought and implemented.

[13] The 15 villages in the household survey were selected on the basis of certain criteria, including the extent of forest cover, product diversity, and year and purpose of establishment of the community forestry program.

[14] Randomization of treatments within a village was not feasible since the experiments had to be rolled out over the course of several days and thus mixing treatments within a village would risk that subjects learn about this variation.

experiment.[15] The enumerators were trained to carefully explain and demonstrate the structure of the experiment according to a script, and they knew the importance of doing so in a neutral manner.[16] A total of 360 subjects participated in the experiment, 180 in each treatment.

3. Results

Table 1 presents descriptive statistics of the sample. As can be seen, about 90% of the household heads are males and over 50% are literate. Farming is the predominant activity for the households and the total size of parcels for own cultivation is slightly over 4 acres. When comparing the socio-economic characteristics between the treatments, there is no statistical difference based on non-parametric tests, except that subjects in the TAKE treatment are statistically significantly younger at the 5% level (49.0 years vs. 46.0 years; Mann-Whitney U-test; $p = 0.03$). However, we do not consider this average difference of 3 years to be of any economic significance. Overall, the randomization of subjects into treatments seems to have worked well.

Table 1 about here

Next, we turn to the impact of framing on unconditional contributions. For ease of comparison, we will mostly refer to "contributions" also in the take frame, and by this we mean the *effective contribution*, i.e., $10 - w_i$, where w_i is the amount withdrawn from the public account. The results are presented in Table 2. Subjects in the GIVE treatment contribute on average 5.02 birr out of the endowment of 10 birr, i.e., 50.2%, whereas the average contribution in the TAKE treatment is 40.0%. In standard one-shot linear public goods games, average voluntary contributions (and normally a give frame is used) usually range from 40% to 60% of endowments (see, e.g., Chaudhuri, 2011). The difference between the two treatments is significant at the 1% level (Mann-Whitney U-test; $p < 0.01$). Furthermore, there is a substantial difference across treatments in the share of subjects opting for free riding by contributing nothing and full contribution, respectively. In the GIVE treatment, 2.2% of the subjects contribute nothing and 7.8% contribute their full endowment. In the TAKE treatment, 19.4% opt to free ride and only 1.1%, i.e., two subjects, contribute the whole endowment. At both ends of the spectrum – zero and full contribution – there is a highly significant difference in proportions

---

[15] We established a credible procedure through the involvement of the Kebele leader in the organization of the experiments, and also by running the experiments, which were approved by the national government, through the Ethiopian Development Research Institute.
[16] See the Appendix for a complete transcript of the instructions.

between the treatments (Kruskal-Wallis tests; $p < 0.01$ in both cases). These results are in line with other experiments and thus we replicate the pattern of higher contributions in the GIVE frame using a non-standard subject pool, namely farmers in Ethiopia. Interestingly, the stark difference in proportion of free riders between the frames is remarkably similar to the results in Andreoni (1995).

Table 2 about here

Result 1. *The average contributions are significantly higher and the proportion of free riders is significantly lower in the GIVE frame.*

We implement a minimum level in both frames and the results are presented in Table 3. In the GIVE treatment, the minimum level has a negative effect on contributions: the average contribution decreases from 5.02 birr to 4.46 birr, i.e., by 0.56 birr, which corresponds to 5.6 percentage points in terms of the initial endowment, and this difference is significant at the 1% level (Wilcoxon signed-rank test; $p < 0.01$). Conversely, in the TAKE treatment the average contribution increases from 4.03 birr in the baseline to 4.74 birr after the introduction of the minimum level, which corresponds to a 7.1 percentage point increase (Wilcoxon signed-rank test; $p < 0.01$). The contributions are significantly different between the two treatments at the 10% level (Mann-Whitney U-test; $p = 0.08$).

Table 3 about here

Result 2. *When a minimum level is introduced in the GIVE frame, contributions to the public good decrease significantly; in contrast, when a minimum level is introduced in the TAKE frame, contributions increase significantly. Thus, the framing has a substantial effect on the efficiency of minimum levels.*

We provide more detailed information regarding the effect of the minimum level in Table 4, which is a disaggregation of the results in Table 3. For the baseline contributions, subjects are split into three categories: below, at, and above the minimum level of 2 birr. For contributions in the second stage, i.e., when the minimum level is implemented, subjects are separated into those who precisely match the minimum level of 2 birr, and those who contribute above it. This results in six possible contribution profiles, and thus we can obtain richer information on how different groups of subjects react to the minimum level and how these reactions differ between the two treatments.

The top two rows of Table 4 make it clear that the effect along the extensive margin is much stronger in the TAKE treatment. Here, the minimum level has a direct effect on 22% of the subjects who contributed less than 2 birr in the baseline, compared with only

4% of the subjects in GIVE. Interestingly, however, in the TAKE treatment about half of the subjects who contributed less than 2 birr in the baseline, i.e., 10% of the total number of subjects in this treatment, substantially increase their contribution in the second stage, on average from 0.6% of the endowment in the baseline to 57.1% when the minimum level has been implemented. Hence, they are strongly crowded in.

Looking at the bottom two rows of the table, we can see that the majority of subjects contribute more than 2 birr in the baseline in (89% in GIVE and 74% in TAKE). When the minimum level is introduced, some of these subjects reduce their contribution all the way down to the postulated minimum and this effect is similar across treatments. Still, the majority of subjects in both treatments voluntarily contribute above the imposed minimum level and herein lays the main source of difference between the treatments, since subjects in the GIVE treatment on average decrease their contributions whereas those in the TAKE treatment instead increase their contributions.

Table 4 about here

Using the full sample, we summarize subjects' reactions along the intensive margin by looking at the extent to which the average subject is crowded in, by voluntarily increasing their contribution, or crowded out, by voluntarily decreasing their contribution, when the minimum level is imposed. We calculate the average difference between subjects' baseline contribution and their contribution when the minimum level has been imposed, counting only the contributions that are above 2 birr, which is the minimum level in both cases, but averaging over all subjects in each treatment.[17] Subjects in the GIVE treatment are crowded out by an average of 0.63 birr (Wilcoxon signed-rank test; $p < 0.01$) while subjects in the TAKE treatment are crowded in by an average of 0.31 birr, yet this effect is only significant at the 10% level (Wilcoxon signed-rank tests; $p = 0.08$). These two effects are statistically different from each other at the 1% level (Mann-Whitney U test; $p < 0.01$), which means that subjects act differently along the intensive margin in GIVE and TAKE.

Among the subjects who already contributed above 2 birr in the baseline, 56% are crowded out and 17% are crowded in in the GIVE treatment, whereas 38% are crowded out and 35% are crowded in in the TAKE treatment. The average crowding-out effect in

---

[17] Mechanically, we follow Falk and Kosfeld (2006) and proceed as follows. First, baseline contributions below the minimum level are manually adjusted up to this level. Then we test whether the average difference between MCL and baseline contributions is different from zero using all subjects (the average difference is zero for contribution profiles in rows 1 and 3 of Table 4). A positive difference implies crowding in and a negative crowding out.

GIVE within this subsample is 0.85 birr, which is significant at the 1% level (Wilcoxon signed-rank test; $p < 0.01$). In the TAKE treatment, the crowding-in effect we found when looking at the full sample does not exist within this subsample and instead there is a weak but insignificant crowding-out effect of 0.19 birr (Wilcoxon signed-rank test; $p = 0.49$).

Result 3. *The minimum level significantly crowds out voluntary contributions to the public good in the GIVE frame but weakly crowds in contributions in the TAKE frame.*

The other part of the public goods experiment elicited conditional contributions following Fischbacher et al. (2001). Subjects stated their contribution for each possible average contribution of the other group members and we can thus investigate the impact of framing and minimum levels on subjects' contribution strategies. Figure 1 displays average conditional contributions in the baseline for both treatments. The positive slope suggests that subjects are on average conditional cooperators, i.e., they contribute more when the other group members contribute more. The contribution schedule in the GIVE frame consistently lies above the schedule in the TAKE frame, which is in line with the theoretical prediction since a given outcome, e.g., the others contribute nothing to the public good, reveals less generosity in TAKE than in GIVE and is thus met with comparatively more negative reciprocity.

Figure 1 about here

Figures 2 and 3 illustrate the main effect of introducing the minimum level. Interestingly, the minimum level seems to have little overall effect on subjects' contribution strategies; only in TAKE and at low levels of others' public good contributions is there a marked difference, but the schedules gradually converge.

Figures 2 and 3 about here

Table 5 provides further information. In column 1, individual contributions in the baseline are regressed on others' average contribution, a treatment dummy, and an interaction term between the two.[18] The estimated intercept – own contribution conditioned on the others contributing nothing to the public good – is 2.68 birr in the base category, i.e., the GIVE treatment. The TAKE dummy identifies the difference in constant compared with the GIVE treatment. The difference is 1.01 birr less, which is significant at the 1% level. The interaction term captures the difference in slope

---

[18] We report from OLS regressions in Table 5 but a Tobit model gives similar results.

compared with the GIVE treatment. It is statistically significant (steeper in TAKE) but the point estimate of 0.06 is small in both absolute economic terms and compared with the estimated slope in GIVE, which is 0.52 birr. By and large, the strong framing effect regarding baseline contributions to the public good thus prevails when we assess subjects' contribution strategies by allowing them to condition their contribution on the other group members' contributions.[19] In column 2, the specification is the same but concerns contributions after the minimum level has been implemented. The main difference compared with the regression in column 1 is that both the TAKE dummy and the interaction term are insignificant, which means that there is no difference in contributions between the two treatments. The regression in column 3 is a difference-in-difference specification and captures how the effect of introducing a minimum level compares across the GIVE and TAKE treatments. The MCL dummy and the interaction with the slope variable (Others) capture the effect of the minimum level in the GIVE treatment. The effect on the constant is an insignificant -0.26 birr and the effect on the slope is significant but not very large. The point estimate on TAKE×MCL captures the difference-in-differences in constant between TAKE and GIVE: when a minimum level is introduced, the effect on the constant is 0.74 birr stronger in TAKE than in GIVE. Interestingly, however, the difference in the effect on the slope goes in the other direction: when the minimum level is introduced, the effect on the slope of the contribution schedule is 0.09 birr weaker in TAKE than in GIVE.

Table 5 about here

Result 4. *The framing has a substantial impact on subjects' contribution strategies. When subjects can condition their decisions on the other group members' contributions, they contribute significantly more in the baseline GIVE frame than in the baseline TAKE frame. To some extent, the framing also has an effect on the efficiency of minimum levels on subjects' contribution strategies.*

Using the Fischbacher et al. design further enables us to classify subjects into contribution types following the convention introduced by these authors. We thus consider subjects as either conditional cooperators, free riders, hump-shaped

---

[19] We can see in the table that there is a positive effect of others' average contribution on individuals' own contribution to the public good and this indicates that subjects on average are conditional cooperators. Another way to investigate this is to look at the correlation between subjects' unconditional contributions and their beliefs about others' unconditional contributions, which is similar to the information we get from the contribution table except for the additional uncertainty that subjects face concerning the accuracy of their beliefs. Regressing unconditional contributions on beliefs yields a positive effect that is significant at the 1% level in three of the four treatment combinations (it is insignificant in the TAKE treatment with a minimum level).

contributors, or others. We define types as follows: conditional cooperators either have a weakly increasing conditional contribution schedule, or their Spearman's rho (correlation with the others' average contribution) is positive and significant at the 1% level; free riders always contribute the lowest amount allowed; hump-shaped contributors display a positive Spearman correlation up to their highest contribution, whereafter the correlation is negative (both correlations should be significant at the 1% level); others do not fit into any of these categories. The distribution of contribution types is displayed in Tables 6 and 7 and we can see it is stable: the baseline type distribution is not significantly different between the GIVE and the TAKE treatment (Kruskal-Wallis test; $p = 0.70$), and, furthermore, 76% of subjects in the GIVE treatment and 74% in the TAKE treatment are classified as the same type in the baseline as in MCL (observations on the diagonal in either table).[20]

Tables 6 and 7 about here

Result 5. *The distribution of contributor types is stable. Neither across the two frames in the baseline nor within each frame with respect to the implementation of a minimum contribution level do contributor type distributions differ significantly.*

4. Conclusion

Contribution to public goods is an important issue for policy makers. In their role as policy makers, however, they have the possibility to act as choice architects. Two institutional factors they can affect fairly easily are whether to frame public goods in terms of giving or taking, and also whether or not to impose a low compulsory minimum level of contributions. These factors are easier to implement than other institutional features such as monetary punishment and exclusion. Our focus is on the relevance for policy design and we conduct our public goods experiments as a lab-in-the-field experiment in rural Ethiopia. The objective of our paper is to investigate the interaction effects between framing and minimum levels.

We find a strong frame dependency in the efficiency of minimum levels. Overall, we find the highest contributions in the give frame without a minimum level. In the standard public goods experiment, the give frame results in significantly higher contributions than the take frame. When the minimum level is introduced, the

---

[20] Thus, in neither treatment is the distribution significantly altered following the introduction of a minimum contribution level (Kruskal-Wallis tests; $p = 0.70$ in GIVE and $p = 0.18$ in TAKE). However, the MCL type distributions do differ between frames (Kruskal-Wallis test; $p = 0.04$).

contribution levels between the frames are reversed and the levels are significantly higher in the take frame. Cooperation is crowded out in the give frame but crowded in in the take frame.

The paper makes several contributions. First, some real-life cooperation problems might exhibit the aspects of the give frame whereas other problems might instead resemble the take frame. For example, the problem of labor contribution to a local public good can be considered a give frame, while the problem of over-harvesting from a local public good can be considered a take frame. Our results indicate that introducing minimum levels in order to curb underinvestment in local public goods is an efficient policy only in the latter case, despite the fact that individuals face the same underlying economic problem in both frames. Furthermore, this finding highlights the potential importance of framing lab experiments in accordance with the structural aspects of the real-life cooperation problem that we as researchers are trying to address. This is a question of external validity and could be relevant to the literature that examines institution formation, policy interventions, and the general impact of incentives on cooperation in public goods experiments, since these results almost exclusively rely on experiments using the give frame (rather than the take frame). Second, changing the framing can sometimes be a policy tool in itself. Our results indicate that the success of a policy intervention, like the introduction of minimum levels in public goods provision, can depend on the manner in which the situation is framed (e.g., whether cooperative behavior is encouraged by focusing on aspects of doing good or avoiding harm). To the extent it is possible, policy makers in the role as choice architect should thus use the best frame in a given context and for a given policy intervention.

Framing and minimum levels are cheap and easy-to-implement local-level policy options, and they could be especially important in situations with comparatively weak centralized formal institutions. However, a potential drawback with a minimum level is the substantial heterogeneity in subjects' reactions to it. It is important to acknowledge the fact that the high proportion of free riders found in experiments using students as subjects has not been replicated to the same extent using non-student samples, making the case for pure efficiency gain of a minimum level weaker but a negative effect of crowding out more possible. We find a high share of conditional cooperators, which policy makers need to account for in their choice of policy instruments. For example, Rustagi et al. (2010) and Kosfeld and Rustagi (2015) implement novel experiments in rural Ethiopia and document the importance of conditional cooperation for successful management of forest commons. By using the design in Fischbacher et al. (2001), our study together with those two papers and the large-scale study by Fosgaard et al. (2014)

contribute to improve our knowledge about people's cooperative preferences in non-student subject pools.

There are potentially large gains in using simple policy regimes, especially in developing countries. Framing and minimum levels in public good provision are two such options. We analyze their interaction and impact on public good provision, looking at the net effects of implementing these institutions. The behavioral effects in our experiments are likely shaped by a combination of several mechanisms and an interesting avenue for future research would be to pin them down in more detail.

Acknowledgments

References

Andreoni, J., 1993. An experimental test of the public-goods crowding-out hypothesis. American Economic Review 83 (5), 317–327.

Andreoni, J., 1995. Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. Quarterly Journal of Economics 110 (1), 1–21.

Bowles, S., Polanía-Reyes, S., 2012. Economic incentives and social preferences: substitutes or complements? Journal of Economic Literature 50 (2), 368–425.

Brown, J., Kling, J., Mullainathan, S., Wrobel, M., 2008. Why don't people insure late-life consumption? A framing explanation of the under-annuitization puzzle. American Economic Review Papers and Proceedings 98 (2), 304–309.

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

Cookson, R., 2000. Framing effects in public goods experiments. Experimental Economics 3 (1), 55–79.

Cox, C., 2015. Decomposing the effects of negative framing in linear public goods games. Economics Letters 126, 63–65.

Cox, J., Friedman, D., Sadiraj, V., 2008. Revealed altruism. Econometrica 76 (1), 31–69.

Cox, J., Ostrom, E., Sadiraj, V., Walker, J., 2013. Provision versus appropriation in symmetric and asymmetric social dilemmas. Southern Economic Journal 79 (3), 496–512.

Cubitt, R.P., Drouvelis, M., Gächter, S., Kabalin, R., 2011. Moral judgments in social dilemmas: how bad is free riding? Journal of Public Economics 95 (3–4): 253–264.

Deci, E.L., 1971. Effects of externally mediated rewards on intrinsic motivation. Journal of Personality and Social Psychology 18 (1), 105–115.

Deci, E.L., 1975. Intrinsic Motivation. Plenum Press, New York.

Dufwenberg, M., Gächter, S., Hennig-Schmidt, H., 2011. The framing of games and the psychology of play. Games and Economic Behavior 73, 459–478.

Eckel, C.C., Grossman, P.J., Johnston R.M., 2005. An experimental test of the crowding out hypothesis. Journal of Public Economics 89 (8), 1543–1560.

Falk, A., Kosfeld, M., 2006. The hidden costs of control. American Economic Review 96 (5), 1611–1630.

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

Fischbacher, U., Gächter, S., 2010. Social preferences, beliefs, and the dynamics of free riding in public good experiments. American Economic Review 100, 541–556.

Fischbacher U., Gächter S., Quercia S., 2012. The behavioral validity of the strategy method in public good experiments. Journal of Economic Psychology 33(4), 897–913.

Fosgaard, T., Hansen, L.G., Wengström, E., 2014. Understanding the nature of cooperation variability. Journal of Public Economics 120, 134–143.

Fryer, R.G., Levitt, S., List, J., Sadoff, S., 2012. Enhancing the efficacy of teacher incentives through loss aversion: a field experiment. NBER Working Paper No. 18237.

Gromet, D., Kunreuther, H., Larrick, R., 2013. Political ideology affects energy-efficiency attitudes and choices. Proceedings of the National Academy of Sciences 110 (23), 9314–9319.

Gronberg, T.J., Luccasen R.A. III, Turocy, T.L., Van Huyck, J.J., 2012. Ate tax-financed contributions to a public good completely crowded-out? Empirical evidence. Journal of Public Economics 96 (7–8), 596–603.

Gächter, S., 2007. Conditional cooperation: behavioral regularities from the lab and the field and their policy implications, in: Frey, B.S., Stutzer, A. (Eds), Economics and Psychology. A Promising New Cross-Disciplinary Field, MIT Press, Cambridge, MA, pp. 19–50.

Gächter, S., Kölle, F., Quercia, S., 2014. The ABC of cooperation in voluntary contribution and common pool extraction games. Working paper.

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

Herrmann, B., Thöni, C., 2009. Measuring conditional cooperation: A replication study in Russia. Experimental Economics 12, 87–92.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., 2001. In search or homo economicus: behavioral experiments in 15 small-scale societies. American Economic Review Papers and Proceedings 91, 73–78.

Henrich, J., Heine, S.J., Norenzayan, A., 2010. The weirdest people in the world? Behavioral and Brain Sciences 33 (2–3), 61–135.

Hossain, T., List, J., 2012. The behavioralist visits the factory: increasing productivity using simple framing manipulations. Management Science 58 (12), 2151–2167.

Khadjavi, M., Lange, A., 2015. Doing good or doing harm: experimental evidence on giving and taking in public goods games. Experimental Economics 18 (3), 432–441.

Kocher, M., Cherry, T., Kroll, S., Netzer, R., Sutter, M., 2008. Conditional cooperation on three continents. Economics Letters 101, 175–178.

Kocher, M, Martinsson, P., Persson, E., Wang, X., 2016. Is there a hidden cost of imposing a minimum contribution level for public good contributions? Working paper.

Kosfeld, M., Rustagi, D., 2015. Leader punishment and cooperation in groups: experimental field evidence from commons management in Ethiopia. American Economic Review 105 (2), 747–783.

Krupka, E.L., Weber, R. 2013. Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary? Journal of the European Economic Association 11(3), 495-524.

Martinsson, P., Pham-Khanh, N., Villegas-Palacio, C., 2013. Conditional cooperation and disclosure in developing countries. Journal of Economic Psychology 34: 148–155.

Martinsson, P., Villegas, C., Wollbrant, C., 2015. Cooperation and social classes: evidence from Colombia. Social Choice and Welfare, forthcoming.

Olken, S., Singhal, M., 2011. Informal taxation. American Economic Journal: Applied Economics 3 (4), 1–28.

Ostrom, E., 1991. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press, Cambridge, UK. Thaler, R., Sunstein, C.R., 2003. Libertarian Paternalism. The American Economic Review 93, 175–179.

Rustagi, D., Engel, S., Kosfeld, M., 2010. Conditional cooperation and costly monitoring explain success in forest common management. Science 330, 961–965

Thaler, R.H., Sunstein, C.R., 2009. Nudge: improving decisions about health, wealth and happiness. Penguin Books, London.

Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. Science 211, 453–4 58.

Vieider, F.M., Lefebvre, M., Bouchouicha, R., Chmura, T., Hakimov, R., Krawczyk, M., Martinsson, P., 2015. Common components of risk and uncertainty attitudes across contexts and domains: evidence from 30 countries. Journal of the European Economic Association 13 (3), 421–452

Ziegelmeyer, A., Schmelz, K., Ploner, M., 2012. Hidden costs of control: four repetitions and an extension. Experimental Economics 15 (2), 323–340.

Appendix A: Instructions for the GIVE treatment

You will now take part in another decision making experiment in which you can earn real money. The experiment consists of two parts that are completely independent of each other. This means that decisions that you and others make in part 1 have no impact on part 2. After completion of part 1, we will explain part 2. The experiment is designed so that your earnings will depend on your decisions and on the decisions of other people who make a similar decision.

You will be a member of a group of four household heads in this village. This means there are three other household heads in your group. We selected your group partners randomly, and you will not know who they are. Neither will the other household heads in your group know who you are. Your group partners will be asked to make exactly the same decision as you. Your earnings will be paid in cash on (*insert date*).

(Repeat)

I will now explain the decision you and each member of your group have to make. Remember that you will earn money based on your decisions. There is no right or wrong decision. You can decide whatever you want.

At the beginning of the experiment, each group member, including you, will receive **10 birr**. Each person has to decide how to divide their **10 birr** into two places: a **private bag** and a **public bag** (show public bag). This means you will make this decision with the 10 birr you will receive (put money on table). You can put **anything between 0 birr and 10 birr** in the public bag and then put the rest in your private bag.

After everyone in your group has made a decision, the payments to you and the other three household heads in your group will be determined as follows:

The amount that you put in the private bag will be saved for you as it is.

The amount that you and your group partners put in the public bag will be counted. Then we will double the total amount in the public bag. This doubled amount will be distributed **equally** among you and your group members regardless of how much everyone put in the public bag in the first place. For example, if the sum of what you and the other three household heads in your group put is 16 birr, we will add another 16 birr to make it 32 birr. Then each one of you will earn 8 birr from the public bag (*show with money*).

The total amount of money that you earn from this experiment is therefore what you put in your private bag, plus what you earn from the doubled amount of what you and your group members put in the public bag, which is shared equally.

(*Repeat*)

Please note that you can decide to put whatever you want in the public bag. The same applies to the other three household heads in your group. They will not know how much you put in the public bag before they decide. And you will not know how much they put in the public bag before you decide. But they too know that whatever is collected in the

public bag will be doubled by us and shared equally, regardless of the amount that each member put in the public bag.

(Repeat)

Now I will ask you some questions to make sure you have understood the decision you will make and how payments are determined.

1. Assume that all group members (including you) put 0 birr into the public bag.
   a. How much will you earn? _____
   b. How much will the other people in your group each earn? _____

2. Assume also that all group members (including you) put the entire 10 birr into the public bag.
   a. How much will you earn? _____

3. Assume that one of the other members in your group puts 5 birr in the public bag, another puts in 3 birr, and the third puts in 2 birr. This means there is 10 birr in the public bag from your group partners.
   a. What is the total you will earn if you put 0 birr in the public bag?

      _____
   b. What is the total you will earn if you put 6 Birr in the public bag?

      _____
   c. What is the total you will earn if you put 10 birr in the public bag?

      _____

Do you have any questions about how payments are calculated? Please ask me if anything is unclear.

You will make two types of decisions.

First, you will choose how much of your 10 birr you want to put in the public bag without knowing what the other three people in your group will decide. Remember that they have to make exactly the same decision as you. **Second**, I will ask you how much of your 10 birr you want to put in the public bag for each possible average contribution of the other group members (rounded to the next integer). Your contribution decision will therefore depend on how much other group members contribute. This will become clear to you when you see the contribution table *(show table)*. If, for example, the other three group members have put 0 birr in the public bag, how much do you want to put in? If the others on average have put 1 birr in the public bag, how much do you want to put in? *(demonstrate)*

Average contribution of other group members *(explain average carefully)*	Your contribution *(write answer)*
0 birr	
1 birr	
2 birr	

3 birr	
4 birr	
5 birr	
6 birr	
7 birr	
8 birr	
9 birr	
10 birr	

After all group members in the experiment have made all decisions, one member of the group will be selected randomly by the supervisor. For this randomly determined group member, only the second decision will be relevant. For the three group members who are not selected by the random mechanism, the first decision will be the relevant decision. When you make your two decisions, you will not know whether you will be randomly selected. You will therefore have to think carefully about both decisions, since either one of them can become relevant for you.

Your payment will either depend on the first or the second decision, and this will be determined by a lottery. First we will draw a lottery to select one person from your group. This person will be paid based on the second decision, while the other three will be paid based on their first decision. Thus, three people in your group will be paid based on how much they contribute without knowing what the other group members contribute. The first decisions of these three people will be used to calculate the average of their contributions. For the fourth individual, the second decision will then be used to determine how much he/she will be paid. For example, if the others on average contributed 4 birr, then the fourth person's contribution will be…. *(point table)*, while if the others on average contributed 8 birr, then the fourth person's contribution will be *(point table).* Because it is a lottery, you could end up being the person who is paid based on the first decision or the second decision. You will therefore have to think carefully about each decision you make.

(Repeat)

Please note that you can decide to put whatever amount you want in the public bag. The same applies to the other three household heads in your group. They will not know how much you put in the public bag before they decide, and you will not know how much they put in the public bag before you decide. But they too will know that whatever is collected in the public bag will be doubled by us and shared equally, regardless of the amount that each member put in the public bag.

(Make sure the person has understood the game before recording decisions. Repeat the whole thing again and again if needed)

Decision 1: How much of your 10 birr do you want to contribute to the public bag?

Answer:_____

Decision 2: Now I will ask you how much you would like to contribute to the public bag after telling you how much the other three contribute.

(For each row below, ask: "If your group members contribute x birr on average to the public bag, how much would you like to contribute to this bag?" Repeat each question.)

Average contribution of other group members *(explain average carefully)*	Your contribution *(write answer)*
0 birr	
1 birr	
2 birr	
3 birr	
4 birr	
5 birr	
6 birr	
7 birr	
8 birr	
9 birr	
10 birr	

Part 2

The basic decision situation in part 2 is the same as in part 1. Again you will be matched in groups of four people. The members of your group will, however, be different in part 2 from your group members in part 1.

Again, you have an endowment of 10 birr. You can put these in a private bag or the public bag. Each of the four group members has to make two types of decisions as before.

In the first decision, you will decide how many birr you want to put in the public bag. **This time, you are required to make a minimum contribution of 2 birr. Each member in your group has to contribute at least 2 birr, and everyone knows this**.

Your payment will be determined in the same way as in the previous part of the experiment.

You will therefore have to think carefully about both decisions, since either one of them can become relevant for you.

Decision 1: How much of your 10 birr do you want to contribute to the public bag? You must contribute at least 2 birr?

Answer:_____

Decision 2: Now I will ask you how much you would like to contribute to the public bag after telling you how much the other three group members contribute on average.

(For each row below, ask: "If your group members contribute x birr on average to the public bag, how much would you like to contribute to the public bag?" Repeat each question)

Average contribution of other group members *(explain average carefully)*	Your contribution *(write answer)*
2 birr	
3 birr	
4 birr	
5 birr	
6 birr	
7 birr	
8 birr	
9 birr	
10 birr	

Now we are done with the decisions. I would like to ask you some additional questions.

1) Think back to the first part of the experiment where you could give anything from 0 to 10 birr.
   a. Please guess how many birr the other three group members have contributed to the public bag in total. Remember that the total contribution put in the public bag by the other three members can be anything from 0 and 30 birr. If you guess the correct amount, you will get 10 birr. So, how many birr do you think the other three group members have given to the public bag in total?

   Answer: _____ birr

   b. Each member in your group has been asked the same guessing question you just answered. What do you think the other three people guessed? Please estimate the sum of the amounts that the other three group members guessed. The sum of the amounts guessed by the other three group members can be anything from 0 to 90 birr. If you guess the correct amount, you will get 10 birr.

   Answer:_____ birr

2) Think back to the second part of the experiment where you decided how much of the 10 birr to contribute, with a 2 birr minimum.
   a. Please guess how many birr the other three group members have contributed to the public bag in total. Remember that the total amount put in the public bag by the other three group members can be anything from 6 to 30 birr. If you guess the correct amount, you will get 10 birr. So, how many birr do you think the other three group members have given to the public bag in total?

Answer: _____ birr

b. Each member in your group has been asked the same guessing question that you just answered. What do you think the other three people guessed? Please estimate the sum of the amounts that the other three group members guessed about the contribution of you and the other two group members. The sum of the amounts guessed by the other three group members can be anything from 18 to 90 birr. If you guess the correct amount, you will get 10 birr.

Answer:_____ birr

Appendix B: Instructions for the TAKE treatment

You will now take part in another decision making experiment in which you can earn real money. The experiment consists of two parts. The two parts are completely independent of each other. This means that decisions that you and others make in part 1 have no impact on part 2. After completion of part 1, we will explain part 2. The experiment is designed so that your earnings will depend on your decisions and on the decisions of other people who make a similar decision.

You will be a member of a group of four household heads in this village. This means there are three other household heads in your group. We selected your group partners randomly, and you will not know who they are. Neither will the other household heads in your group know who you are. Your group partners will be asked to make exactly the same decision as you. Your earnings will be paid in cash on *(insert date)*.

(Repeat)

I will now explain the decision you and each member of your group have to make. Remember that you will earn money based on your decisions. There is no right or wrong decision. You can decide whatever you want.

You and the other three people in your group have to decide what to do with a sum of money that we have put in a **public bag**. There is 40 birr in the public bag. Each one of you can withdraw up to one-fourth or 10 birr from the public bag. Each person will only know how much he or she withdraws and not how much other group members withdraw. This means the other three people in your group will not know how much you choose to withdraw from the public bag. Remember that the maximum allowed for each person is 10 birr.

After everyone in your group has decided how much to **withdraw** *from the public bag, we will count what is left and* **double** *it*. This amount will then be distributed **equally** among you and your group members regardless of how much everyone withdrew from the public bag. For example, if the sum of what is left in the public bag is 16 birr, we will add another 16 birr to make it 32 birr. Then each one of you will earn 8 birr from the public bag (*show with money*).

So, the total money you will earn from this experiment is what you initially withdraw from the public bag plus what you get after the remaining sum is doubled and shared equally.

(Repeat).

Please note that you can decide to withdraw whatever you want from the public bag. The same applies to the other three household heads in your group. They will not know how much you withdraw from the public bag before they decide. And you will not know how much they withdraw from the public bag before you decide. But they too know that whatever remains in the public bag will be doubled by us and shared equally, regardless of the amount that each member withdrew from the public bag.

(Repeat)

Now I will ask you some questions to make sure you have understood the decision you will make and how payments are determined.

1. Assume that all group members (including you) withdraw 10 birr from the public bag.
   a. How much will you earn? _____
   b. How much will the other people in your group earn each? _____

2. Assume that all group members (including you) withdraw 0 birr from the public bag.
   a. How much will you earn? _____

3. Assume that one of the other members in your group withdraws 5 birr from the public bag, another withdraws 3 birr, and the third withdraws 8 birr. This means your group members withdraw a total of 16 birr.
   a. What is the total you will earn if you withdraw 0 birr from the public bag? _____
   b. What is the total you will earn if you withdraw 4 birr from the public bag? _____
   c. What is the total you will earn if you withdraw 10 birr from the public bag? _____

Do you have any questions about how payments are calculated? Please ask me if anything is unclear.

You will make two types of decisions.

First, you will choose how much you want to withdraw from the public bag without knowing what the other three people in your group will decide. Remember that they have to make exactly the same decision as you. **Second**, I will ask you how much you want to withdraw from the public bag for each possible average withdrawal by the other group members (rounded to the next integer). Your withdrawal decision will therefore depend on how much other group members withdraw. This will become clear to you when you see the withdrawal table *(show table)*. If, for example, the other three group members have withdrawn 0 birr each from the public bag, how much do you want to withdraw? If the others on average have withdrawn 1 birr from the public bag, how much do you want to withdraw? *(demonstrate)*

Average withdrawal by other group members *(explain average carefully)*	Your withdrawal *(write answer)*
0 birr	
1 birr	
2 birr	
3 birr	
4 birr	
5 birr	

6 birr	
7 birr	
8 birr	
9 birr	
10 birr	

After all group members in the experiment have made all decisions, one member of the group will be selected randomly by the supervisor. For this randomly determined group member, only the second decision will be relevant. For the three group members who are not selected by the random mechanism, the first decision will be the relevant decision. When you make your two decisions, you will not know whether you will be randomly selected. You will therefore have to think carefully about both decisions, since either one of them can become relevant for you.

Your payment will either depend on the first or the second decision, and this will be determined by a lottery. First we will draw a lottery to select on person from your group. This person will be paid based on the second decision, while the other three will be paid based on their first decision. Thus, three people in your group will be paid based on how much they withdraw without knowing what the other group members have withdrawn. The first decisions of these three people will be used to calculate the average of their withdrawal. For the fourth individual, the second decision will then be used to determine how much he/she will be paid. For example, if the others on average have withdrawn 4 birr, then the fourth person's withdrawal will be…. *(point table)*, while if the others on average have withdrawn 8 birr, then the withdrawal by the fourth individual will be *(point table).* Because it is a lottery, you could end up being the person who is paid based on the first decision or second decision. You will therefore have to think carefully about each decision you make.

(Repeat)

Please note that you can decide to withdraw whatever amount you want from the public bag. The same applies to the other three household heads in your group. They will not know how much you withdraw from the public bag before they decide, and you will not know how much they withdraw from the public bag before you decide. But they too know that whatever is left in the public bag will be doubled by us and shared equally, regardless of the amount that each member withdrew from the public bag.

(Make sure the person has understood the game before recording decisions. Repeat the whole thing again and again if needed)

Decision 1: How much do you want to withdraw from the public bag?

Answer:_____

Decision 2: Now I will ask you how much you would like to withdraw from the public bag after telling you how much the other three have withdrawn on average.

(For each row below, ask: "If your group members withdraw x birr on average from the public bag, how much would you like to withdraw from this bag?" Repeat each question)

Average withdrawn by other group members *(explain average carefully)*	Your withdrawal *(write answer)*
0 birr	
1 birr	
2 birr	
3 birr	
4 birr	
5 birr	
6 birr	
7 birr	
8 birr	
9 birr	
10 birr	

Part 2

The basic decision situation in part 2 is the same as in part 1. Again you will be matched in groups of four people. The members of your group will, however, not be the same as in part 1.

Each of the four group members, including you, has to make two types of decisions as before.
In the first decision, you will decide how much you want to withdraw from the public bag, which has 40 birr in it. **This time, you are required to withdraw a maximum of 8 birr, or leave at least 2 birr in the public bag. Each group member has to leave at least 2 birr in the public bag**.

Your payment will be determined in the same way as in the previous part of the experiment.

You will therefore have to think carefully about both decisions, since either one of them can become relevant for you.

Decision 1: How much do want to withdraw from the bag, given that you have to leave at least 2 birr?

Answer:_____

Decision 2: Now I will ask you how much you would like to withdraw from the public bag after telling you how much the other three withdraw on average.

(For each row below, ask: "If your group members withdraw x birr on average from the public bag, how much would you like to withdraw from the public bag?" Repeat each question)

Average withdrawal by other group members *(explain average carefully)*	Your withdrawal *(write answer)*
0 birr	
1 birr	
2 birr	
3 birr	
4 birr	
5 birr	
6 birr	
7 birr	
8 birr	

Now we are done with the decisions. I would like to ask you some additional questions.

1. Think back to the first part of the experiment where you could withdraw any amount from 0 to 10 birr from the public bag.
   a. Please guess how many birr the other three group members have withdrawn from the public bag in total. Remember that the total they can withdraw from the public bag can be anything from 0 to 30 birr. If you guess the correct amount, you will get 10 birr. So, how many birr do you think the other three group members have withdrawn from the public bag in total?

   Answer: _____ birr

   b. Each member in your group has been asked the same guessing question that you just answered. What do you think the other three people guessed? Please estimate the sum of the amounts that the other three group members guessed. The sum of the other three group members' guessed amounts can be anything from 0 to 90 birr. If you guess the correct amount, you will get 10 birr.

   Answer:_____ birr

2. Think back to the second part of the experiment where you could withdraw from the public bag but had to leave at least 2 birr from the maximum 10 birr you could have withdrawn.
   a. Please guess how many birr the other three group members have withdrawn from the public bag in total. Remember that the total amount they can withdraw from the public bag can be anything from 0 to 24 birr. If you guess the correct amount, you will get 10 birr. So, how many birr do you think the other three group members have withdrawn from the public bag in total?

Answer: _____ **birr**

b. Each member in your group has been asked the same guessing question that you just answered. What do you think the other three people guessed? Please estimate the sum of the amounts that the other three group members guessed about the withdrawal by you and the other two group members. The sum of the other three group members' guessed amounts can be anything from 0 to 72 birr. If you guess the correct amount, you will get 10 birr.

Answer:_____ **birr**

Tables

Table 1. Descriptive statistics of subjects (household heads) separated by treatment.

	GIVE	TAKE	H_0: No difference between GIVE and TAKE (*p-value*[*])
Age	49.0	46.0	0.03
Male	87.2%	90%	0.40
Can read and write	51.1%	50.6%	0.92
Farming is main activity	91%	91%	0.86
Trust Kebele	2.51	2.46	0.56
Off-farm labor	23%	28.3%	0.23
Household size	7.0	6.8	0.40
Parcel size (acres)	4.56	4.17	0.38
Has savings	19.4%	23.3%	0.37
Remittances	7.8%	10.6%	0.36
Observations	180	180	

Note: [*]Mann-Whitney U test. *Trust Kebele:* Kebele is the smallest administrative unit in Ethiopia and the question is phrased, "Most people who live in your Kebele can be trusted" (1 = agree fully, 5 = disagree fully); *Off-farm labor:* at least one household member worked off the farm at least once during last year; *Parcel size:* the size of the household's land used for own cultivation; *Remittances:* the household received a remittance during last year.

Table 2. Effect of framing on unconditional contributions in the baseline.

	Average	Contributes nothing	Contributes everything
GIVE (*n* = 180)	5.02 (2.15)	2.2%	7.8%
TAKE (*n* = 180)	4.00 (2.44)	19.4%	1.1%
p-value	< 0.01[a]	< 0.01[b]	< 0.01[b]

Note: Standard deviations in parentheses. *n* = number of subjects. [a] Mann-Whitney U test. [b] Kruskal-Wallis test.

Table 3. Effect of a minimum level on unconditional contributions (separated by treatment).

	Baseline	MCL	*Difference*	*p-value*
GIVE (*n* = 180)	5.02 (2.15)	4.46 (1.98)	-0.56	< 0.01[b]
TAKE (*n* = 177)	4.03 (2.44)	4.74 (2.06)	+0.71	< 0.01[b]
p-value	< 0.01[a]	0.08[a]	< 0.01[a]	

Note: Standard deviations in parentheses. *n* = number of subjects. [a] Mann-Whitney U test. [b] Wilcoxon signed-rank test (H_0: *Difference* = 0). Three subjects in the TAKE treatment report unfeasible contributions in MCL and we thus exclude them from the analysis.

Table 4. Disaggregation of the effect on unconditional contributions.

Contribution profile	GIVE			TAKE		
(Baseline, MCL)	Share of subjects	Baseline contr.	MCL contr.	Share of subjects	Baseline contr.	MCL contr.
(<2, =2)	(2%)	0.33	2.00	(12%)	0.14	2.00
(<2, >2)	(2%)	0.50	3.75	(10%)	0.06	5.71
(=2, =2)	(4%)	2.00	2.00	(2%)	2.00	2.00
(=2, >2)	(3%)	2.00	4.50	(2%)	2.00	6.00
(>2, =2)	(8%)	4.90	2.00	(10%)	5.82	2.00
(>2, >2)	(81%)	5.60	4.90	(64%)	5.21	5.56

Note: The share-of-subjects columns show the frequency of subjects who contributed according to the profile described in column one. The cells show the average contribution to the public good in the baseline and in MCL among subjects who behave according to this pattern, for GIVE and TAKE, respectively.

Table 5. Effect of framing and minimum level on cooperative preferences (OLS regression).

Dependent variable: conditional contribution	Baseline (1)	MCL (2)	Full sample (3)
Constant	2.68 (0.16)***	2.45 (0.15)***	2.70 (0.17)***
Others	0.52 (0.03)***	0.58 (0.03)***	0.53 (0.03)***
TAKE	-1.01 (0.25)***	-0.30 (0.24)	-1.02 (0.26)***
TAKE $\times$ Others	0.06 (0.04)***	-0.03 (0.04)	0.06 (0.04)
MCL			-0.26 (0.16)
MCL $\times$ Others			0.05 (0.02)**
TAKE $\times$MCL			0.74 (0.22)***
TAKE $\times$MCL $\times$ Others			-0.09 (0.03)**
Observations	3916	3096	6800
- Clusters	356	344	340

Note: Robust standard errors clustered on individuals (in brackets). *Others* is the average contribution of the others in the group. *TAKE* is a treatment dummy and *MCL* is a dummy for the second stage public goods game (when the minimum level is imposed). The dependent variable is individual i's conditional contribution. *** denotes significance at the 1% level, ** significance at the 5% level, and * significance at the 10% level. Four subjects either did not fill out the contribution table or reported unfeasible contributions in the baseline and these subjects are excluded from the analysis in (1). A further sixteen subjects made a similar mistake after the introduction of the minimum level and thus sixteen and twenty subjects are excluded from the analysis in (2) and (3), respectively.

Table 6. Distribution of contribution types in GIVE ($n = 168$).

Baseline	MCL	Conditional cooperators	Min-level ("free riders") (=2)	Hump-shaped	Others	Total
Conditional cooperators		56.55%	1.19%	0.00%	9.52%	67.26%
Free riders (=0)		0.00%	0.60%	0.00%	0.00%	0.60%
Hump-shaped		0.00%	0.00%	0.00%	1.19%	1.19%
Others		0.60%	0.60%	1.79%	19.05%	30.95%
Total		66.07%	2.38%	1.79%	29.76%	100%

Note: Type definition follows Fischbacher et al. (2001). Conditional cooperators: increasing schedule or positive Spearman's rho (correlation with others' average contribution) at 1%; Free riders: contribute the lowest amount allowed; Hump-shaped: positive (negative) Spearman correlation at 1% up to (beyond) their highest contribution. Twelve subjects in this treatment report unfeasible conditional contributions after the introduction of the minimum level and are therefore excluded from the analysis.

Table 7. Distribution of contribution types in TAKE ($n = 172$).

Baseline	MCL	Conditional cooperators	Min-level ("free riders") (=2)	Hump-shaped	Others	Total
Conditional cooperators		47.09%	0.00%	0.00%	13.37%	60.47%
Free riders (=0)		1.16%	1.74%	0.00%	0.00%	2.91%
Hump-shaped		1.74%	0.00%	0.00%	2.33%	4.07%
Others		5.23%	0.58%	1.16%	25.58%	32.56%
Total		55.23%	2.33%	1.16%	41.28%	100%

Note: Type definition follows Fischbacher et al. (2001). Conditional cooperators: increasing schedule or positive Spearman's rho (correlation with others' average contribution) at 1%; Free riders: contribute the lowest amount allowed; Hump-shaped: positive (negative) Spearman correlation at 1% up to (beyond) their highest contribution. Eight subjects in this treatment report unfeasible conditional contributions after the introduction of the minimum level and are therefore excluded from the analysis.
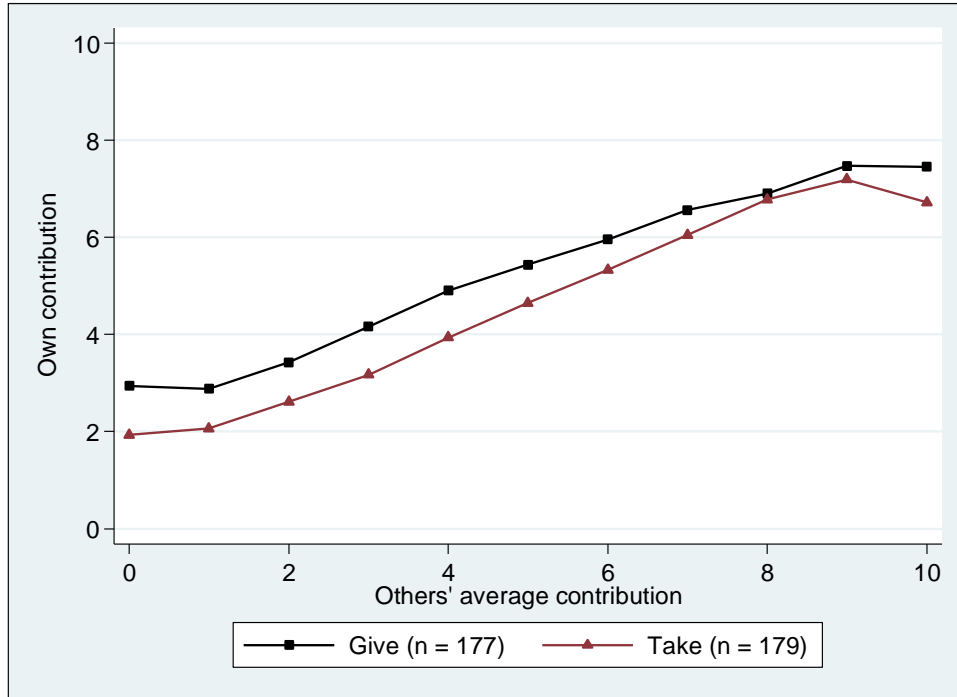
Figures

Figure 1. Average conditional contributions in the baseline (separated by treatment).

Note: One subject in the GIVE treatment and three subjects in the TAKE treatment either did not complete the contribution table or reported unfeasible numbers. These four subjects are excluded from the analysis.

Figure 2. Average conditional contributions in GIVE ($n = 168$).

Note: Twelve subjects in this treatment reported unfeasible conditional contributions after the introduction of the minimum level and are therefore excluded from the analysis.

Figure 3. Average conditional contributions in TAKE ($n = 172$).
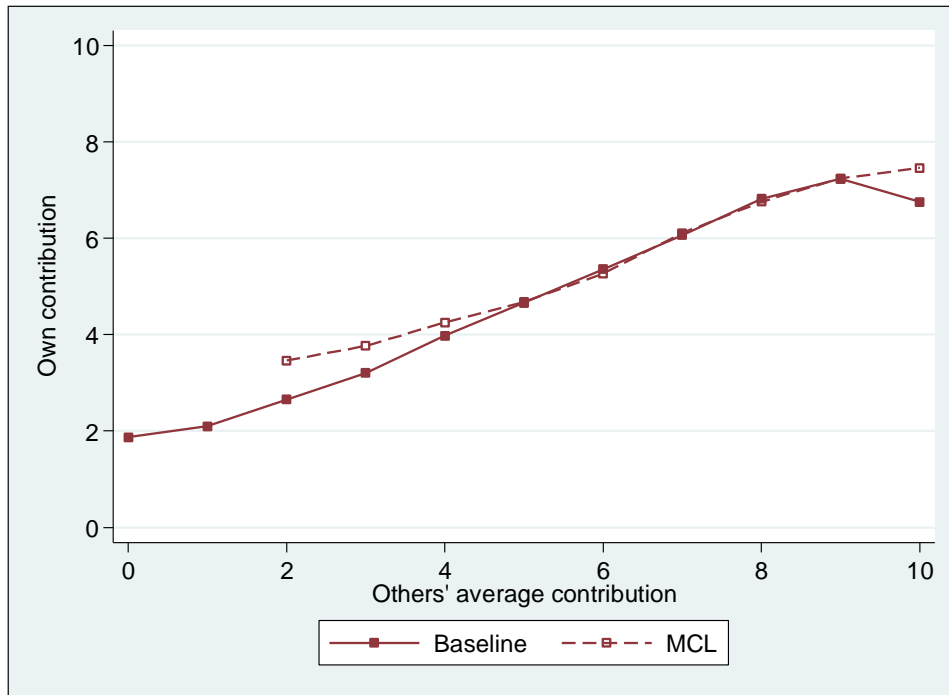
Note: Eight subjects in this treatment reported unfeasible conditional contributions after the introduction of the minimum level and are therefore excluded from the analysis.
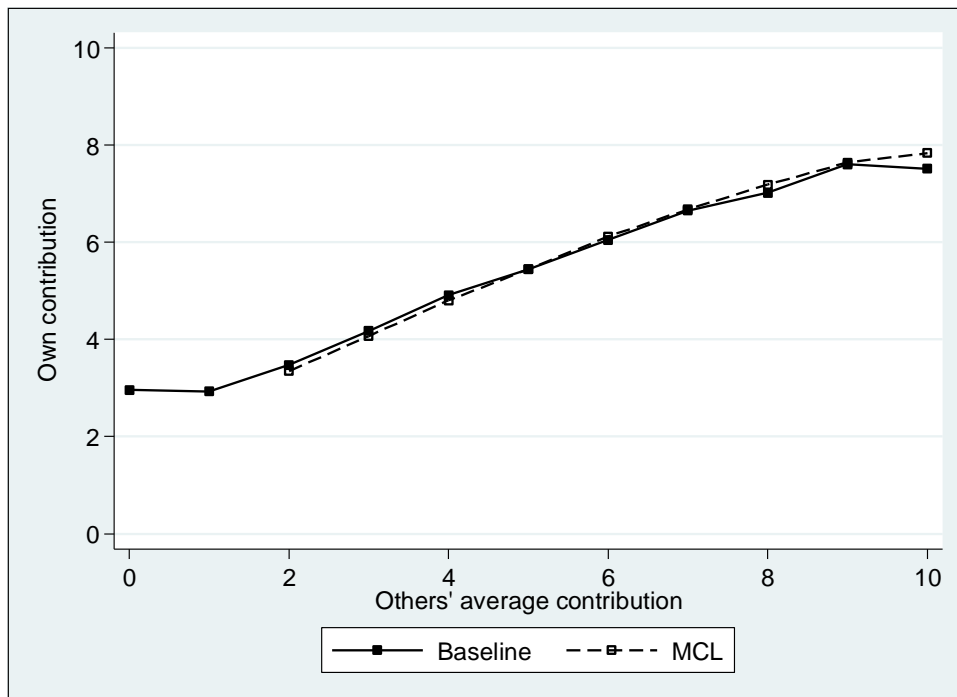
Paper III

Frustration and Anger in Games: A First Empirical Test of the Theory

Emil Persson*

Department of Economics, University of Gothenburg, Sweden

Abstract

Anger can be a strong behavioral force, with important consequences for human interaction. For example, angry individuals may become hostile in their dealings with others, and this has strategic consequences. Battigalli, Dufwenberg, and Smith (2015; BDS) develop a formal framework where frustration and anger affect interaction and shape economic outcomes. This paper designs an experiment testing the predictions based on central concepts of their theory. The focus is on situations where other-responsibility is weak or nonexistent, and in this specific context I find only limited support for the theory: While unfulfilled expectations about material payoffs seem to generate negative emotions in subjects, which is in line with BDS' conceptualization of frustration, behavior is generally not affected by these emotions to the extent predicted by the theory.

Keywords: Emotion, Anger, Blame, Psychological games, Experiment.
JEL Codes: C72, C91, D03.

---

*Department of Economics, University of Gothenburg, Box 640, SE 405 30 Gothenburg, Sweden; Tel.: +46 31 786 2927; E-mail: emil.persson@economics.gu.se.

1   Introduction

Anger can be a strong behavioral force, with important consequences for human interaction. For example, angry individuals may become hostile in their dealings with others, which could shape interaction and outcomes in, for example, situations involving negotiation and bargaining, contractual holdup, delegated decision making, conflict, and social dilemmas.[1]

Although it seems important to understand the sources of anger, as well as its consequences for strategic interaction, this topic has received relatively little attention in the development of behavioral theory.[2] Battigalli, Dufwenberg, and Smith (2015; henceforth, BDS) contribute to fill this gap in the literature by developing a theory where frustration and anger affect interaction and outcomes, using the framework of psychological game theory (Geanakoplos et al., 1989; Battigalli and Dufwenberg, 2009). The objective of my paper is to investigate the empirical relevance of BDS, and to this end I design an experiment that tests the predictions based on central concepts of their theory. This constitutes the first empirical test of BDS.

In BDS, anger is anchored in frustration, which is the result of unfulfilled expectations about material payoffs.[3] Frustration sometimes makes players hostile toward their co-players. When frustrated, a player may go after other players, but his desire to do so depends on his evaluation of the other players' part in the outcome that frustrates him. BDS develop three different versions of how this evaluation process shapes the actions of frustrated players: With *simple anger*, frustrated players are angry with anyone, regardless of the source of frustration; with *anger from blaming behavior*, players are targeted only if they caused frustration; and with *anger from blaming intentions*, players are targeted only if they intended to cause frustration.

With this paper I develop empirical tests for two key features of the theory: simple anger and anger from blaming behavior. Simple anger (BDS' first anger hypothesis) formalizes a version of the classical Dollard et al. frustration-aggression-displacement hypothesis, where aggression through a displacement effect is directed at substitute targets (Dollard et al., 1939; Berkowitz, 1989). Such displaced aggression could be relevant when the source of frustration is intangible, as in the case of an unexpected loss suffered by a local soccer team, which has been associated with substantial increases in domestic violence and violent crime (Card and Dahl, 2011; Munyo and Rossi, 2013). Dramatic

---

[1] For empirical accounts of negative emotions and anger in similar situations, see e.g., Pillutla and Murnighan (1996), Sanfey et al. (2003), Bosman and van Winden (2002), Bolle et al. (2014), Hopfensitz and Reuben (2009), and Drouvelis and Grosskopf (2014).

[2] This is especially true for the analysis of immediate (as opposed to anticipated) emotions, which focuses on the "action tendency" of emotions experienced by the decision maker (e.g., Elster, 1998; Loewenstein, 2000).

[3] This is based on the notion of frustration as an obstruction to reaching a desired outcome, which is a common conceptualization in psychology; see, e.g., BDS or Potegal and Stemmler (2010) for details and discussion.

changes in weather and climate could have a similar effect, for instance through strong and correlated income shocks as a source of frustration and anger.[4] With anger from blaming behavior (BDS' second anger hypothesis), frustrated individuals are angry with those who caused frustration through their behavior. Vis-à-vis simple anger, a key ingredient for anger from blaming behavior is the importance of blame attribution and other-responsibility;[5] and vis-à-vis anger from blaming intentions, anger from blaming behavior is more relevant if people focus more on what they can observe (and expect to observe) rather than on intentions, which in many situations are more difficult to discern.

BDS create a rich framework for theoretical analysis. It is, however, quite challenging to develop convincing empirical tests of the theory, partly because people's behavior is shaped both by the emotions they experience themselves and by their anticipation of others' behavior due to the emotions they might experience (it is a psychological game where subjects have belief-dependent utility functions), and partly also because the different versions of the theory may predict similar behavioral patterns in a given situation of interest. Therefore, the empirical strategy in the present paper is to lift up, focus on, and compare key features of the theory in a condensed and specific setting, as simply as possible. The experimental treatments are built around the following situation: A player who had a good chance to earn 100 Swedish kronor (about 12 US dollars at the time of writing) finds himself with only 10 kronor. Is he frustrated? Theory suggests he might be, since he has been obstructed from reaching a desired outcome. Would he punish a passive co-player, who had no chance at all to prevent the misfortune? Simple anger suggests he might. Would he punish an active co-player, who made a "bad" choice in a binary lottery and thereby caused the misfortune? Anger from blaming behavior suggests he might.

I develop specific empirical tests for simple anger and anger from blaming behavior. The tests are quite extreme in that they focus on relatively unsophisticated behavior, such as Pareto-damaging punishment of passive co-players. Theoretically, simple anger is widely applicable to situations involving sophisticated strategic interaction, but the conceptual basis is quite rudimentary: I hit my head on the kitchen shelf and therefore I punish you; I hit my thumb with a hammer and therefore I punish you; or you bring me bad news and therefore I punish you (Frijda, 1993). Focusing on this aspect seems natural for a first empirical test of the theory. Moreover, while anger from blaming behavior admits more sophisticated reasoning (about co-players' blameworthiness), it too can be quite rudimentary. For example, it admits punishment for mistakes or bad luck. Gurdal et al. (2013) document behavior that is consistent with this aspect of anger from blaming behavior. In their experiment, an agent invests money on behalf of a

---

[4]See, e.g., Burke et al. (2015) and Ranson (2014) for a discussion and empirical evidence on climate and interpersonal violence.

[5]See, e.g., Averill (1983), Smith and Ellsworth (1985), and Wranik and Scherer (2010) for a discussion on the role of blame attribution in anger.

principal. The agent chooses between a safe and a risky prospect, and the principal subsequently decides on remuneration for the agent and a dummy player. Interestingly and in line with anger from blaming behavior, Gurdal et al. find that agents are paid less (relative to the dummy player) following bad realizations of the risky prospect, i.e., they are punished for bad uncontrollable luck.[6] Building on these insights, I specifically test whether aspects of blame, frustration, and punishment in BDS are relevant in situations where other-responsibly is weak or nonexistent. The focus of the paper is thus on emotion and behavior at one end of a spectrum, where there is complete payoff symmetry between players (i.e., no economic inequality) and where frustration is caused by chance moves or bad luck rather than harmful actions with blameworthy intent.

Results from the experiment indicate that unfulfilled expectations about material payoffs generate negative emotions in subjects, which is in line with BDS' conceptualization of frustration. However, these negative emotions do not generally affect subjects' behavior to the extent predicted by the theory. The experiment focuses on situations where other-responsibility is weak or nonexistent, and in this specific context I do not find much support for the simple anger hypothesis or the anger from blaming behavior hypothesis.

The rest of the paper is organized as follows. Section 2 describes relevant aspects of BDS' theory and illustrates the concepts of simple anger and anger from blaming behavior using two examples. Section 3 describes and motivates the experimental design, which is based on operationalized versions of these two examples, and derives hypotheses for the experiment. The results are presented in Section 4, and in the final section I discuss the findings and conclude the paper.

2 Frustration and anger

This section describes BDS' theory (short version) using two examples to illustrate the concepts of *simple anger* (Section 2.1) and *anger from blaming behavior* (Section 2.2). Later on, these examples will be operationalized and implemented as the two main treatments in the experiment (Section 3).

There are two players and two stages. Actions are represented by a sequence $(a^t)_{t=1,2}$, where $a^t = (a^t_i, a^t_j)$ and $i \neq j$. For instance, a^1 represents the actions taken in the first stage such that $h = a^1$ is a history of length one, and a^2_i represents player i's action in the second stage. Conditional on a history of actions, i's belief about his own action (which

---

[6]On punishment for mistakes or bad luck, see also Cushman et al. (2009), Gino et al. (2010), and Bartling and Fischbacher (2012); but cf. Rand et al. (2015). Furthermore, Mollerstrom et al. (2015) and Cappelen et al. (2015) find that irrelevant or forced decisions affect how third-party spectators are willing to compensate people for bad luck. In particular, Cappelen et al. (2015) implement an experiment where subjects are allocated money based on a procedure that is very similar to the one used in the present paper, and they find strong effects on redistribution decisions ex post. There is also a related literature on "outcome bias," which examines how negative uncontrollable outcomes affect how other persons are evaluated (e.g., Baron and Hershey, 1988; Tan and Lipe, 1997).

is interpreted as his plan) and i's belief about j's action are given by α_i. The material end-game payoff for i depends on actions taken during the course of the game (a^1 and a^2) and is denoted π_i.

Anger is anchored in frustration, which is a key aspect of the theory. It is the result of unfulfilled expectations, i.e., when the best attainable outcome falls short of a person's expectation ex ante (which means that he was obstructed from reaching a desired outcome). At $t = 2$, i's frustration is given by

$$F_i(a^1; \alpha_i) = \left[\mathbb{E}\left[\pi_i; \alpha_i\right] - \max_{a_i^2} \mathbb{E}\left[\pi_i \mid (a^1, a_i^2); \alpha_i\right] \right]^+ , \tag{1}$$

where $[x]^+ = \max\{x, 0\}$. That is, it is the difference between the end-game payoff i expected at $t = 1$ (first term) and the maximal end-game payoff i realizes that he can secure for himself at $t = 2$ (second term). Frustration is thus the result of unexpected bad outcomes beyond i's control.

Frustration sometimes makes players hostile toward their co-player. When i is frustrated, he may go after j, but his desire to do so depends on his evaluation of j's part in the outcome that frustrates him. BDS develop three different versions of how this evaluation process shapes the actions of frustrated players: With *simple anger* (SA), player i is angry with j regardless of the source of frustration; in contrast, with *anger from blaming behavior* (ABB), player j will be targeted only if he caused i's frustration, and with *anger from blaming intentions* (ABI), player j will be targeted only if his intention was to frustrate i. The behavior of player i moving at history h is determined by the following general utility function:

$$U_i(h, a_i; \alpha_i) = \mathbb{E}\left[\pi_i \mid (h, a_i); \alpha_i\right] - \theta_i B_{ij}(h; \alpha_i)\mathbb{E}\left[\pi_j \mid (h, a_i); \alpha_i\right], \tag{2}$$

where $\theta_i \geq 0$ is a sensitivity parameter. When i is frustrated and under the impression it is j's fault (which is governed by the blame function $B_{ij} \in [0, F_i]$), he will consider forgoing own material payoff (π_i) in order to reduce j's payoff (π_j). The formula in (2) thus implements punishment as the action tendency of frustration. The three versions of anger are separated by the blame function $B_{ij}(h; \alpha_i)$, which determines what amount of frustration i attributes to j, i.e., to what extent i blames j for his frustration. This function is always less than or equal to i's frustration, $F_i(h; \alpha_i)$.

For simple anger, blame always equals frustration:

$$B_{ij}^{SA}(h; \alpha_i) = F_i(h; \alpha_i). \tag{3}$$

For anger from blaming behavior, i's blame of j depends on j's behavior. BDS consider two different functional forms of the blame function for anger from blaming

behavior: With *could-have-been blame*, player i considers what "could have been" (at most) had j chosen a different action, and with *blaming unexpected deviations*, player i considers instead the material payoff he would have obtained had j behaved as i expected him to. The blame function for the former, could-have-been blame, is:

$$B_{ij}^{ABB}(h; \alpha_i) = \min \left\{ \left[\max_{h' \prec h, a_j' \in A_j(h')} \mathbb{E}\left[\pi_i \mid (h', a_j'); \alpha_i\right] - \mathbb{E}\left[\pi_i \mid h; \alpha_i\right] \right]^+, F_i(h; \alpha_i) \right\}, \quad (4)$$

thus capturing the difference between the maximal expected payoff j could have given i through his action a_j' (choosing from all possible actions A_j) at history h' (which precedes h) and the payoff that i currently (at h) expects to secure for himself at the end of the game. Turning to the latter functional form of the blame function for anger from blaming behavior, blaming unexpected deviations, the blame function is instead:

$$B_{ij}^{ABB}(h; \alpha_i) = \min \left\{ \left[\sum_{h' \prec h, a_j' \in A_j(h')} \alpha_{ij}(a_j') \mathbb{E}\left[\pi_i \mid (h', a_j'); \alpha_i\right] - \mathbb{E}\left[\pi_i \mid h; \alpha_i\right] \right]^+, F_i(h; \alpha_i) \right\},$$
$$(5)$$

such that i considers his expected payoff had j behaved as i expected ($a_{ij}(a_j')$ is the marginal probability with which i thinks that j chooses action a_j'), and compares it with his current expected payoff (second term).[7]

2.1 A game with a correlated income shock

BDS illustrate the difference between blame modulated through simple anger and anger from blaming behavior (and anger from blaming intentions) with an example (Figure 1). A carpenter (Andy) and his co-worker (passive player Bob) are at work. On a good day (G), they both earn a high wage of 2 and thus everything is fine. However, on the occasional bad day (B) Andy accidentally hits his thumb with a hammer. He becomes frustrated and may turn on his co-worker Bob (T) even though he did not cause the upset, or, alternatively, he accepts things as they are (N) and they can move on with their production. Chance determines whether B or G prevails, with probability p and $(1 - p)$, respectively.

The bad day thus comes with a correlated income shock, which frustrates Andy if the maximal payoff at $t = 2$ (which is 1) is less than the payoff Andy expected at $t = 1$. Frustration thus depends on p and on Andy's belief at $t = 1$ about his own action in the event of B, i.e., how he plans to react when frustrated due to B. In formal terms, Andy

---

[7]With BDS' third version of blame, anger from blaming intentions, i takes j's intentions to frustrate i into account: he compares the material payoff resulting from j's action with the material payoff he thinks that j thought would result from his action (blame thus depends on i's belief about j's belief about i's action). I don't test this aspect of BDS' theory and thus the interested reader is referred to BDS for formal treatment and further discussion.
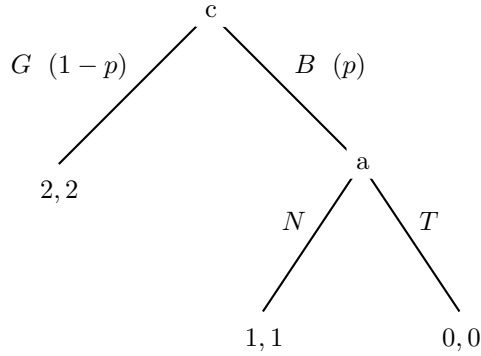
Figure 1: The simple anger game ("Hammering one's thumb")

initially expected to earn a material payoff of $\mathbb{E}[\pi_a; \alpha_a] = 2(1-p) + \alpha_a(N|B)p$, where $\alpha_a(N|B) \in [0,1]$ is Andy's plan for the bad day B, i.e., the probability with which he expects to choose N should B occur. On the bad day B, the maximal payoff Andy can secure for himself (by choosing $a_a^2 = N$) is 1. Thus, using (1), his frustration at $t = 2$ is given by:

$$F_a(B; \alpha_a) = 2(1-p) + \alpha_a(N|B)p - 1 \tag{6}$$

and it is increasing in $\alpha_a(N|B)$ and also in $-p$, since this increases the payoff Andy initially expects.

With simple anger, Andy might punish Bob for the frustrating outcome B. This is Pareto damaging in terms of material payoffs but still preferable to Andy if he is sufficiently "prone to anger" (which is determined by the sensitivity parameter θ_a in the utility function). Using (2), Andy's utility from N is $1 - \theta_a[2(1-p) + \alpha_a(N|B)p - 1] \cdot 1$, and his utility from T is 0. Thus, Andy will punish Bob (by choosing T) if θ_a is sufficiently large, since then $U_a(T) \geq U_a(N)$. In contrast, neither with anger from blaming behavior nor with anger from blaming intentions does Andy blame Bob for B (so $B_{ab} = 0$), because Bob had no possibility to affect the outcome in the game. Hence, Andy will not punish Bob under these two versions of the theory.

2.1.1 Equilibrium

BDS employ the notion of sequential equilibrium (SE) adapted to psychological games by Battigalli and Dufwenberg (2009).[8] The SE concept requires that action profiles are sequentially rational, such that no gains can be made by deviating from planned actions, and that beliefs are correct. In the game between Andy and Bob above, there are three

---

[8]They explore a second solution concept, "polymorphic sequential equilibrium," in order to develop a different conceptualization of players' intentions. A discussion of this concept is therefore beyond the scope of this paper.

types of equilibria when frustration is modulated through simple anger (and $p < 1/2$):[9]

(i) If θ_a is large enough, Andy cares a lot about frustration and thus he will punish Bob after B ($a_a^2 = T$). There is an SE with $\alpha_a(N|B) = 0$ and $U_a(B,T;\alpha_a) \geq U_a(B,N;\alpha_a)$, such that $\theta_a \geq \frac{1}{1-2p}$.

(ii) If θ_a is small enough, Andy does not care much about frustration and thus he will not punish Bob after B ($a_a^2 = N$). There is an SE with $\alpha_a(N|B) = 1$ and $U_a(B,T;\alpha_a) \leq U_a(B,N;\alpha_a)$, such that $\theta_a \leq \frac{1}{1-p}$.

(iii) For a small intermediate interval of θ_a, Andy's belief that he will punish Bob after B is increasing in θ_a (i.e, in the extent to which he cares about frustration). There is an SE with $\alpha_a(N|B) = \frac{1}{p\theta_a} - \frac{1-2p}{p}$ and $U_a(B,T;\alpha_a) = U_a(B,N;\alpha_a)$, such that $\theta_a \in (\frac{1}{1-p}, \frac{1}{1-2p})$.

2.2 A game with a correlated income shock and some blame

In order to bring anger from blaming behavior into play and make it comparable with simple anger, consider the following modification of the simple anger game displayed in Figure 1: Chance moves L or R with probability $(1-p)$ and p, respectively, and thereafter Bob chooses between l and r without having observed the chance move. Bob's choice leads to states identical to B and G in the simple anger game, but whether l or r is the "good" choice that leads to G depends on chance: l leads to G and r to B after L, and vice versa after R. The game is displayed in Figure 2.
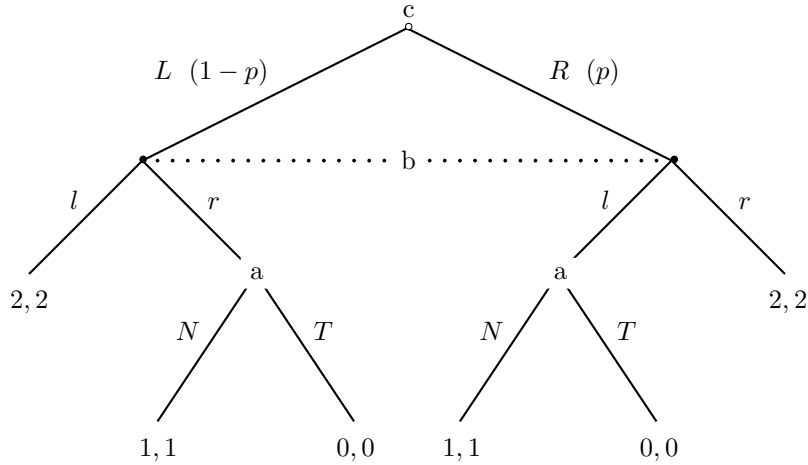

Figure 2: The blame-behavior game

For ease of comparison with the simple anger game, we restrict attention to situations where $p < 1/2$. Andy's frustration depends on his belief (at the root of the game)

---

[9]We consider $p < 1/2$ since it may generate frustration. (When $p \geq 1/2$, Andy will not be frustrated after B and thus he will always choose N in this case.)

about Bob's plan, and we let $\alpha_a(l)$ denote the probability with which Andy expects Bob to choose l. If $\alpha_a(l) = 1$, Andy expects l with certainty and his plan for the bad day B is given by $\alpha_a(N|(R,l))$, i.e., the probability with which he expects to choose N in the event of B. Andy's frustration on the bad day is thus given by $F_a((R,l); \alpha_a) = 2(1-p) + \alpha_a(N|(R,l))p - 1$ and it is similar to frustration in the simple anger game, given by (6) above. The equation says that if Andy expects l with certainty, his frustration after B is increasing in the probability of L, i.e. the probability with which l leads to the good day G since this increases his ex-ante expected payoff. With simple anger, there are then equilibria where $\alpha_a(l) = 1$ and, as in the simple anger game above, (i) Andy with $\theta_a \geq \frac{1}{1-2p}$ punishes Bob after B, (ii) Andy with $\theta_a \leq \frac{1}{1-p}$ does not punish Bob after B, and (iii) Andy with $\theta_a \in (\frac{1}{1-p}, \frac{1}{1-2p})$ is indifferent between punishing and not punishing Bob after B.[10]

Bob moves before Andy and this opens for the possibility of Andy blaming and punishing Bob in accordance with anger from blaming behavior. With could-have-been blame (first version of anger from blaming behavior), predictions coincide with simple anger predictions. The reason is that Andy blames Bob for frustrating him, since whenever the bad day occurs it would have been avoided had Bob chosen differently. For example, if the bad day is due to (R,l), it would have been avoided had Bob chosen r instead of l. Even though Bob from an ex-ante perspective could not prevent the income shock (beyond choosing the action that maximizes their chance of the good day G, which is l when $p < 1/2$), by (4) Andy fully blames Bob for his frustration. In contrast, with blaming unexpected deviations (second version of anger from blaming behavior), by (5) Andy would not blame Bob and therefore not punish him after (R,l) if l is what he expected at the root of the game, and vice versa for (L,r) if r is what he initially expected.[11] This illustrates the difference between could-have-been blame and blaming unexpected deviations, since only with the former is there an SE where Bob is punished in the blame-behavior game.

In general, Andy is as frustrated on the bad day in the simple anger game (Figure 1) as in the blame-behavior game (Figure 2). Since Bob is always blamed for the bad day in the latter game using could-have-been blame, comparing across games we have exogenous variation in the extent to which Bob can be blamed for a given level of frustration. Simple anger is relevant in both games but anger from blaming behavior is relevant in only one of them, and this constitutes the basic idea for the empirical strategy used in the present paper.

---

[10]For Bob, l is optimal when $p < 1/2$ since it maximizes the chance of the good day G with the high payoff (and, given Andy's plan, Bob deviating would not affect Andy's aggression on the bad day B).

[11]Of course, if Andy initially expected r but Bob deviated by choosing l, Andy would blame Bob for the bad day occurring due to (R,l) since it would have been avoided had Bob behaved as Andy expected. This is the basis for an SE where Andy expects r and Bob chooses r in order to deflect blame, even though l would maximize the chance of the good day G when $p < 1/2$.

3 Experiment

3.1 Design

The main objective of this paper is to test condensed versions of simple anger and anger from blaming behavior (could-have-been blame). To this end, I operationalize the simple anger game (Figure 1) and the blame-behavior game (Figure 2) in two separate between-subject treatments: a *Simple* treatment (Figure 3, p. 11) and a *Blame-behavior* treatment (Figure 4, p. 12).

Henceforth, Andy is called player **a** and Bob player **b**. Two important modifications are made when operationalizing the games: First, since frustration in the bad state B originates in a low p, I increase the outside payoff and set $p = 1/2$ in order to be able to implement the games using the direct-response method rather than the strategy method (with a low p the direct-response method would consume an unreasonable amount of observations). Therefore, frustration in the experimental treatments originates in a high outside payoff rather than in a low p. The choice of p at exactly $1/2$ makes it as easy as possible for subjects to understand and calculate their expected and counterfactual payoffs in each game. It also makes **a**'s frustration in the blame-behavior game independent of his belief about **b**'s behavior, since the bad state B is equally likely whether he chooses l or r (and conversely for the good state G). This removes a potential source of bias in the comparison across the treatments, since otherwise we would have to elicit and condition on **a**'s belief about **b**'s behavior in the *Blame-behavior* treatment. A second modification is that the choice set of **a** is expanded to allow for gradual punishment decisions. It is an implication of BDS that a sufficiently frustrated player i will punish j as much as possible if he blames j for his frustration, and this is due to the linearity of the utility function in (2). Interestingly, Bosman and van Winden (2002) emphasize this aspect of emotional punishment in the context of their power-to-take game, and with the modified choice set of **a** the experimental treatments admit investigation of this aspect in the context of BDS' theory.[12]

A third treatment, *Simple-strategy*, implements the *Simple* treatment using the strategy method, such that **a** decides on punishment in the event of the bad state, i.e., before knowing whether B or G obtains (if B occurs the choice is implemented). An interesting aspect of BDS is that preferences are own-plan dependent. In the simple anger game (Figure 1), this means that **a**'s decision utility (2) at $t = 2$ depends on how he planned at $t = 1$ that he would choose at $t = 2$. That is, **a**'s decision utility $t = 2$ depends on the probability with which he expects (at $t = 1$) to choose N should B occur. If **a**

---

[12]In the power-to-take game, the first mover can appropriate part of the responder's income but the responder can retaliate by destroying some or all of his or her income. There seems to be a strong link between negative emotions and appropriation and destruction behavior in this game, and also between unfulfilled expectations and punishment, which is in line with BDS' conceptualization of frustration (Bosman and van Winden, 2002; Ben-Shakhar et al., 2007).
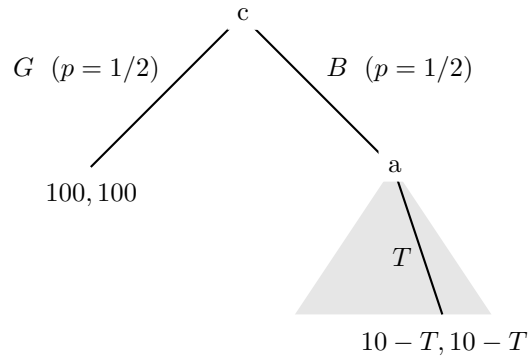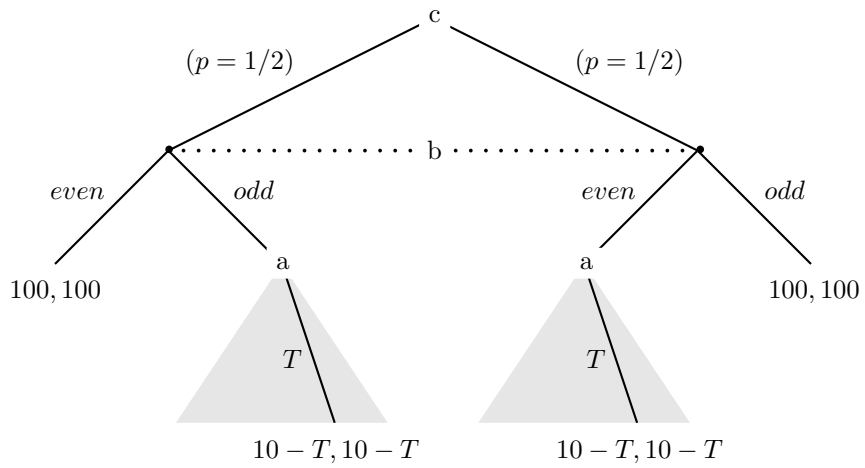
Figure 3: The *Simple* treatment

Figure 4: The *Blame-behavior* treatment

plans to choose N, he becomes more frustrated after B (since he plans to avoid costly punishment and thus expects a higher material payoff) and thus more likely to choose T than had he instead planned to choose T from the beginning. Thus, there is a potential conflict between **a**'s plan at $t = 1$ and his action at $t = 2$, since at $t = 1$ he is not frustrated and thus wants to plan for N, which maximizes his material end-game payoff, but at $t = 2$ he is frustrated and thus wants to punish **b** if sufficiently prone to anger. In equilibrium, **a** takes this plan dependence into account, making a plan that is consistent with him maximizing utility function (2) when frustrated due to B.[13] In light of this, the *Simple-strategy* treatment variation tests whether **a** at the root of the game (in the *Simple-strategy* treatment) seems to have the same goal as **a** has after the frustrating outcome B (in the *Simple* treatment).[14]

In a fourth and final treatment, *Control*, subjects participate in an experiment consisting only of the subgame after B in the *Simple* treatment. The test of the theory (simple anger) is not particularly strong based only on the *Simple* treatment since it rests on rejecting a point prediction at the corner of the choice set, i.e., no punishment. One worry is that some subjects could be influenced by experimenter-demand effects, especially since the game is rather special. With a between-subjects comparison across *Simple* and *Control*, this potential confound is avoided and the test of the theory becomes stronger.

3.2   Procedural details

Subjects are welcomed to the lab and randomly assigned a cubicle workstation. They are given plenty of time to read the instructions (see Appendix) and ask questions (in private) before the beginning of the experiment. They are matched in pairs but decisions are made in private and they never learn the identity of the person they are matched with.

The end-game payoffs in the treatments are as follows (see also Figures 3 and 4): In the good situation G, **a** and **b** receive 100 Swedish kronor each (around 12 US dollars), and in the bad situation B, they receive 10 Swedish kronor each and **a** decides whether to accept the low payoff or punish **b** at a 1:1 cost. Thus, $(\pi_a, \pi_b) = (100, 100)$ in G, and $(\pi_a, \pi_b) = (10 - T, 10 - T)$ with **a** choosing $T \in \{0, 1, \ldots, 10\}$ in B.

In the *Simple* treatment, the public toss of a six-sided die determines whether the good situation G (with the high income) or bad situation B (with the low income) obtains for any pair of subjects. Each subject-pair is given a group number at the beginning of the

---

[13]This can for example be illustrated with the third SE mentioned on p. 8, where **a** plans for N after B (with probability $\alpha_a(N \mid B) = \frac{1}{p\theta_a} - \frac{1-2p}{p}$) so as to keep himself indifferent between N and T when frustrated due to B.

[14]The between-subject test is in the spirit of BDS' multi-self approach, with one "self" of i moving at each history h of the game. See BDS and Battigalli and Dufwenberg (2009, pp. 28) for details.

experiment, and if there is an even-even or odd-odd match between the number shown on the die and the group number, the subject-pair is assigned the good situation (with the high income); in case of an even-odd or odd-even mismatch, the subject-pair is assigned the bad situation (with the low income). For example, a subject-pair with group number 2 will be assigned the good situation if the die shows 2, 4, or 6, and the bad situation for numbers 1, 3, and 5. In contrast, in the *Blame-behavior* treatment, **b** will guess whether the die will show an even or an odd number and if the guess is correct (an even-even or odd-odd match with the number on the die), the good situation is assigned, and if the guess is wrong the bad situation is assigned. Beyond this, everything is identical across the two treatments. In the *Simple-strategy* treatment, subjects first make their decision (in the event of the bad situation) and then the die toss determines what situation will be payoff relevant. In the *Control* treatment, subjects immediately face a situation identical to the low-income situation in the other three treatments.[15]

In the bad situation and in the *Control* treatment, **b** is asked for her belief about **a**'s decision. If correct, **b** is monetarily rewarded. As a measure of frustration, subjects are asked in a post-experimental questionnaire to describe the intensity of emotions felt when learning whether they would end up in the good situation (with the high income) or in the bad situation (with the low income).[16] They answered this question approximately five minutes after they had learned what outcome would be relevant for them.[17] The experiment was computerized with z-Tree (Fischbacher, 2007) and participants were recruited using ORSEE (Greiner, 2015). Sessions lasted around 30 minutes and subjects were on average paid 107 Swedish kronor (around 12.50 US dollars at the time of the experiment).

3.3   Hypotheses

The first hypothesis concerns frustration. By (1), **a** is frustrated when the bad state B obtains since he realizes that his maximal payoff is less than the payoff he expected at the root of the game. Therefore, I investigate whether the correlated income shock generates negative emotions in subjects; i.e., whether those who end up in the bad situation with the low income in the *Simple*, *Blame-behavior*, and *Simple-strategy* treatments on average experience a higher intensity of negative emotions than subjects in the good situation (with the high income):

---

[15]In order to keep the expected hourly wage as close as possible across the treatments, subjects in the *Control* treatment were paid for answering an additional survey at the end of the session.

[16]I follow Bosman and van Winden (2002) and let subjects report on a range of emotions and not just the few (negative) emotions of interest in this study. For a discussion on self-reports as an emotion elicitation method, see, e.g., Robinson and Clore (2002).

[17]The exact formulation read as follows: "We will now ask you to describe how you felt at the moment **when** you found out which situation would be assigned your group (i.e., right after the experimenter had tossed the die). Please indicate the **intensity of each of the emotions** listed below. Please note that there is no 'right' or 'wrong' answer."

Hypothesis 1. *Frustration*: Subjects who end up in the bad situation (with the low income) will report a stronger intensity of negative emotions than subjects who end up in the good situation (with the high income).

The second hypothesis concerns simple anger, and thus the comparison between the *Simple* treatment and the *Control* treatment. In the *Simple* treatment, simple anger predicts full punishment ($T = 10$) if **a** is sufficiently prone to anger ($\theta_a \geq \frac{1}{40}$) whereas anger from blaming behavior predicts zero punishment.[18] In the *Control* treatment, both simple anger and anger from blaming behavior predict zero punishment since there is no frustration.

Hypothesis 2. *Simple anger*: A larger fraction of players **a** will punish their co-player **b** in the *Simple* treatment than in the *Control* treatment.

The third hypothesis concerns anger from blaming behavior (could-have-been blame) and thus comparisons across the *Blame-behavior*, *Simple*, and *Control* treatments. In the *Blame-behavior* treatment, simple anger and anger from blaming behavior predictions coincide (full punishment if sufficiently prone to anger; $T = 10$),[19] whereas only simple anger predicts punishment in the *Simple* treatment, and neither simple anger nor anger from blaming behavior predicts punishment in the *Control* treatment.

Hypothesis 3. *Anger from blaming behavior*: A larger fraction of players **a** will punish their co-player **b** in the *Blame-behavior* treatment than in the *Simple* treatment, AND the same fraction of players **a** will punish their co-player **b** in the *Simple* treatment and the *Control* treatment.

The fourth hypothesis concerns **a**'s plan (at the root of the game) to punish **b**, should the frustrating state B occur. Preferences are own-plan dependent and this creates a potential conflict between **a**'s plan at the root of the game and his action when potentially frustrated due to B. In equilibrium, **a** takes this plan dependence into account (as described on p. 11) and we expect players **a** at the root of the game in the *Simple-strategy* treatment to have the same goal as **a** has after the frustrating outcome B in the *Simple* treatment:

Hypothesis 4. *Consistent planning*: The same fraction of players **a** will punish their co-player **b** in the *Simple* treatment and the *Simple-strategy* treatment.

---

[18]There is a simple-anger SE where **a** chooses full punishment at $t = 2$ (and at $t = 1$ he is certain that he will do so), since by (1) his frustration is $F_a = 0.5 \cdot 100 + 0.5 \cdot 0 - 10 = 40$ and thus by (2) we have $U(T = 0) \leq U(T = 10)$ when $10 - \theta_a \cdot 40 \cdot 10 \leq 0$, i.e. when $\theta_a \geq \frac{1}{40}$.

[19]They are identical to the simple-anger prediction in the *Simple* treatment (punish if $\theta_a \geq \frac{1}{40}$), since (i) the level of frustration in the bad state B is the same in both treatments, and (ii) with could-have-been blame **a** fully blames **b** for his frustration, since the good state G would have obtained had **b** made a different choice.

The fifth and final hypothesis concerns **b**'s belief about **a**'s punishment decision, in each of the four treatments. It is an equilibrium requirement in BDS that players are, at least on average, correct in their beliefs about other players' behavior, and this is what we expect to observe in the experiment:

> **Hypothesis 5.** *Belief about others' punishment*: There is no difference between the fraction of players **a** who punish and the fraction of players **b** who expect to be punished in any of the four treatments.

4 Results

A total of 342 students participated in the experiment. Table 1 displays their participation in the different parts of the experiment (good or bad situation), which depends on the chance move and on **b**'s decision in the *Blame-behavior* treatment. We can see that

Table 1: Participation in the different parts of the experiment.

Treatment	Sessions	Participants	Situation Good	Bad	Players **a** who can make a punishment decision
Simple	11	98	46	52	26
Blame-behavior	17	142	84	58	29
Simple-strategy	5	54	26	28	27
Control	7	48	–	–	24
Total	40	342	156	138	106

Note: In *Simple* and *Blame-behavior*, one player **a** for every subject-pair (i.e., one **a** and one **b**) who experience the bad situation B can make a punishment decision. For instance, in *Simple*, 52 subjects experience the bad situation and out of these, 26 are a player **a** and another 26 are a player **b**. In *Simple-strategy*, decisions are made before chance determines whether situation B or G obtains, and thus 27 of the 54 participants will decide as **a**.

52 subjects in *Simple* and 58 subjects in *Blame-behavior* experience the correlated income shock, and thus subsequently make decisions in the bad situation (players **a** decide on punishment and players **b** report their expectation). Additionally, in the *Simple-strategy* treatment, 54 subjects decide in the same situation but before the chance move, and a further 48 subjects make decisions in the *Control* treatment (which is identical except for the income shock). As can be seen in the table, we have 26, 29, 27, and 24 punishment decisions and an equal amount of reported expectations in *Simple*, *Blame-behavior*, *Simple-strategy*, and *Control*, respectively, and we have 156 subjects who end up in the good situation (with the high income) and 138 subjects who instead experience the bad situation (with the low income).

First we will analyze subjects' emotional reactions. In line with Hypothesis 1, we can see in Table 2 that subjects are more angry and more irritated when hit by the

income shock (the bad situation). The strongest impact is on irritation, with an average intensity of 3.10 within the full sample on a scale ranging from 1 (no emotion at all) to 7 (high intensity of emotion). This finding is in line with Bosman and van Winden (2002), who also document substantial irritation (average intensity of 4.05) following a positive take rate in the power-to-take game.[20] This result supports BDS' conceptualization of frustration as unfulfilled expectations about material payoffs.

> **Result 1:** The bad situation, where subjects' maximal income from the experiment is substantially lower than the ex-ante expected income, generates negative emotions.

Table 2: Average intensity of negative emotions (by situation, G or B).

	Anger			Irritation		
	Good	Bad	*P-value*	Good	Bad	*P-value*
Simple ($N = 74$)	1.29	1.95	< 0.01	1.26	2.73	< 0.01
Blame-behavior ($N = 142$)	1.12	2.07	< 0.01	1.23	3.21	< 0.01
Simple-strategy ($N = 54$)	1.04	2.07	< 0.01	1.12	3.43	< 0.01

Note: I use two-sided Mann-Whitney U tests. After the experiment, subjects were asked to describe the intensity of eleven different emotions (both positive and negative) felt when learning whether the situation with the high income (good) or the situation with the low income (bad) would obtain (1 = no emotion at all; 7 = high intensity of emotion). Due to technical problems, 24 subjects could not answer the question in *Simple* and thus $N = 74$.

The correlated income shock thus generates negative emotions and the question is whether this has a significant effect on subjects' actions and beliefs, following Hypotheses 2–5. Table 3 displays the fraction of subjects who punish (**a**) and expect to be punished (**b**), respectively, for each of the four treatments.[21] We can see that the results are similar across the treatments: 15%–17% of players **a** chose to punish their co-player, and 24%–37% of players **b** expected to be punished. The similarity in behavior across all four treatments indicates the absence of treatment effects and we verify this in panel A of Table 4. The results for Hypotheses 2–4 are formulated as follows:

> **Result 2:** Subjects do not seem to act in accordance with the simple anger hypothesis in the context of the experiment. (We cannot reject the null hypothesis of no difference in behavior between the two treatments *Simple* and *Control*.)

---

[20]Calculated from Table 2 in Bosman and van Winden (2002) as the average irritation intensity reported by all 39 responders in the experiment (the take rate was positive for 36 of them).

[21]The raw data for punishment decisions made by all players **a** who participated in the experiment can be found in the Appendix, Table 5.

Result 3: Subjects do not seem to act in accordance with the anger from blaming behavior hypothesis in the context of the experiment. (We cannot reject the null hypothesis of no difference in behavior between the two treatments *Blame-behavior* and *Simple.*)

Result 4: There is no difference between decisions made at the root of the game (about punishment, should the bad situation occur) and decisions made when the bad situation has occurred. (We cannot reject the null hypothesis of no difference in behavior between the two treatments *Simple* and *Simple-strategy.*)

The fifth and final hypothesis concerned subjects' beliefs about others' punishment decisions. This issue is investigated by comparing the fraction of players **a** who punish with the fraction of players **b** who expect to be punished. Comparing across columns in Table 3, we can see that there is a tendency for players **b** to overshoot in their beliefs about the extent to which their matched player **a** will punish them. For instance, in the *Simple* treatment, 35% of players **b** thought that they would be punished after the correlated income shock, but only 15% of players **a** did so. Looking at the bottom panel of Table 4 we can see that the difference between punishment behavior and expected punishment behavior is significant at the 1% level within the full sample (Mann-Whitney U test; $p < 0.01$, $N = 212$).

Result 5: Subjects in the experiment do not seem to hold correct beliefs about other players' punishment decisions.

Taken together, subjects who end up in the bad situation (with the low income) experience negative emotions, but this does not seem to affect their behavior to the extent predicted by the theory. In the context of the experiment, using a minimalistic design and a deliberatively weak form of blame attribution, I do not find much support for the simple anger or the anger from blaming behavior hypothesis.

Table 3: Fraction of subjects who punish and expect to be punished, respectively.

	Punishment (player **a**)		Expected (player **b**)	
	No	Yes	No	Yes
Simple	22 (85%)	4 (15%)	17 (65%)	9 (35%)
Blame-behavior	24 (83%)	5 (17%)	22 (76%)	7 (24%)
Simple-strategy	23 (85%)	4 (15%)	17 (63%)	10 (37%)
Control	20 (83%)	4 (17%)	15 (63%)	9 (37%)

Note: The table displays the fraction of players **a** who chose not to punish ($T = 0$) and punish ($T > 0$), respectively, and the fraction of players **b** who expected their assigned player **a** to choose not to punish and punish, respectively.

Table 4: Testing Hypotheses 2–5.

Panel A: Punishment

Null hypothesis	P-values
Punish in *Simple* = Punish in *Control*	0.90 ($N = 50$)
Punish in *Blame-behavior* = Punish in *Simple*	0.85 ($N = 55$)
Punish in *Simple* = Punish in *Simple-strategy*	0.95 ($N = 53$)

Panel B: Expectations vs. punishment

Null hypothesis	P-values
Punish = Expected, in *Simple*	0.11 ($N = 52$)
Punish = Expected, in *Blame-behavior*	0.52 ($N = 58$)
Punish = Expected, in *Simple-strategy*	0.06 ($N = 54$)
Punish = Expected, in *Control*	0.11 ($N = 48$)
Punish = Expected, in all treatments	< 0.01 ($N = 212$)

Note: I use two-sided Mann-Whitney U tests based on proportions displayed in Table 3.

5 Conclusion

Anger can be a strong behavioral force, and this has important consequences for human interaction. I design an experiment that tests the predictions based on central concepts in a theory of anger developed by Battigalli, Dufwenberg, and Smith (2015; BDS). Specifically, I test condensed versions of their simple anger and anger from blaming behavior hypotheses in two-player settings, and I investigate aspects of beliefs and compare subjects' planned choices at different stages of the game. The experiment implements a strong and correlated income shock and subsequently tests whether it generates negative emotions in subjects, whether it makes them punish their co-players, and whether they punish their co-players when they can be blamed for the correlated income shock.

The results from the experiment show that the income shock generates negative emotions in subjects; it makes them angry and irritated. This is in line with BDS, who model frustration as a result of unfulfilled expectations about material payoffs. However, these negative emotions do not generally affect punishment decisions to the extent predicted by the theory: Subjects who experience the correlated income shock (the bad situation) do not punish their co-player more than subjects who do not experience the income shock, not even when the co-player can be blamed for the income shock. Thus, in this specific context there is not much support for the notions of simple anger and anger from blaming behavior.

The focus of the paper is on situations where other-responsibility is weak or nonexistent, and where there is complete payoff symmetry between the players. Frustration is caused by chance moves or bad luck rather than harmful actions with blameworthy intent. Focusing on these aspects seems natural for a first empirical test of the theory, since it isolates the relatively unsophisticated punishment behavior that is the essence of simple anger and that also seems to be relevant for anger from blaming behavior, as indicated by data from experiments in Gurdal et al. (2013).

The difference between the results in this study and in Gurdal et al. is interesting. In their experiment, an agent invests money on behalf of a principal. The agent chooses between a safe and a risky prospect, and the principal subsequently decides on remuneration for the agent and a dummy player. They find that agents are blamed and punished for bad uncontrollable outcomes, since they are paid less (relative to the dummy player) following bad realizations of the risky prospect. One important contextual difference between Gurdal et al. and the *Blame-behavior* treatment in the present paper is that the agents in Gurdal et al. do have some ex-ante control over the outcome, in their choice between the safe and the risky prospect, even though they are punished for the bad realization of the risky prospect, which they cannot control. This makes the agent's role in the bad outcome more salient and perhaps this is why anger from blaming behavior seems to be more relevant in this context. Having some form of choice, even if irrelevant,

possibly affects how one's responsibility for a bad outcome is evaluated.[22]

Another important aspect of the design is that payoffs are symmetric, which allows for testing the theory in isolation of distributional concerns. However, maybe the influence of frustration and anger would be stronger in contexts with asymmetric payoffs. For example, disastrous income shocks caused by catastrophic events such as extreme weather or natural disasters might well make people angry, but maybe they are more likely to be angry with people who had the means and economic privileges to prevent or limit personal damage rather than with people who also were badly affected by these events. An interesting extension is thus to test the theory in a context where there is payoff inequality, implemented by chance or by other players in the experiment.

Anger and frustration may shape economic outcomes in profound ways. Angry individuals become hostile in their dealings with others and this has strategic consequences. BDS provide a rich framework for theoretical analysis and I have provided a first test of empirical relevance. There are several interesting avenues for future research.

Acknowledgments

---

[22]This is consistent with Mollerstrom et al. (2015) and Cappelen et al. (2015), who find that irrelevant or even forced decisions affect how third-party spectators are willing to compensate people for bad luck.

References

[1] Averill, J.R., 1983. Studies on anger and aggression: implications for theories of emotion. American Psychologist 38 (11), 1145–1160.

[2] Baron, J., Hershey, J.C., 1988. Outcome bias in decision evaluation. Journal of Personality and Social Psychology 54 (4), 569–579.

[3] Bartling, B., Fischbacher, U., 2012. Shifting the blame: on delegation and responsibility. Review of Economic Studies 79 (1), 67–87.

[4] Battigalli, P., Dufwenberg, M., Smith, A., 2015. Frustration and anger in games. IGIER Working Paper 539.

[5] Battigalli, P., Dufwenberg, M., 2009. Dynamic psychological games. Journal of Economic Theory 144 (1), 1–35.

[6] Ben-Shakar, G., Bornstein, G., Hopfensitz, A., van Winden, F., 2007. Reciprocity and emotions in bargaining using physiological and self-report measures. Journal of Economic Psychology 28 (3), 314–323.

[7] Berkowitz, L., 1989. Frustration-aggression hypothesis: examination and reformulation. Psychological Bulletin 106 (1), 59–73.

[8] Bolle, F., Tan, J.H.W., Zizzo, D.J., 2014. Vendettas. American Economic Journal: Microeconomics 6 (2), 93–130.

[9] Bosman, R., van Winden, F., 2002. Emotional hazard in a power-to-take experiment. Economic Journal 112 (476), 147–169.

[10] Burke, M., Hsiang, S.M., Miguel, E., 2015. Climate and conflict. Annual Review of Economics 7, 577–617.

[11] Cappelen, A.W., Fest, S., Sørensen, E.Ø., Tungodden, B., 2015. Choice and personal responsibility: what is a morally relevant choice? Manuscript, NHH Norwegian School of Economics.

[12] Card, D., Dahl, G.B., 2011. Family violence and football: the effect of unexpected emotional cues on violent behavior. Quarterly Journal of Economics 126 (1), 103–143.

[13] Cushman, F., Dreber, A., Wang, Y., Costa, J., 2009. Accidental outcomes guide punishment in a "trembling hand" game. PLoS ONE 4 (8): e6699

[14] Dollard, J., Doob, L., Miller, N., Mowrer, O., Sears, R., 1939. Frustration and Aggression. Yale University Press, New Haven, CT.

[15] Drouvelis, M., Grosskopf, B., 2014. The effects of anger and happiness on pro-social behaviour. Working paper.

[16] Elster, J., 1998. Emotions and economic theory. Journal of Economic Literature 36 (1), 47–74.

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

[18] Frijda, N.H., 1993. The place of appraisal in emotion. Cognition and Emotion 7 (3–4), 357–387.

[19] Geanakoplos, J., Pearce, D., Stacchetti, E., 1989. Psychological games and sequential rationality. Games and Economic Behavior 1 (1), 60–79.

[20] Gino, F., Shu, L.L., Bazerman, M.H., 2010. Nameless + harmless = blameless: when seemingly irrelevant factors influence judgment of (un)ethical behavior. Organizational Behavior and Human Decision Process 111 (2), 93–101.

[21] Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. Journal of the Economic Science Association 1 (1), 114–125.

[22] Gurdal, M.Y., Miller, J.B., Rustichini, A., 2013. Why blame? Journal of Political Economy 121 (6), 1205–1247.

[23] Hopfensitz, A., Reuben, E., 2009. The importance of emotions for the effectiveness of social punishment. Economic Journal 119 (540), 1534–1559.

[24] Loewenstein, G., 2000. Emotions in economic theory and economic behavior. American Economic Review 90 (2), 426–432.

[25] Mollerstrom, J., Reme, B.A., Sørensen, E.Ø, 2015. Luck, choice and responsibility — an experimental study of fairness views. Journal of Public Economics 131, 33–40.

[26] Munyo, I., Rossi, M.A., 2013. Frustration, euphoria, and violent crime. Journal of Economic Behavior and Organization 86, 136–142.

[27] Pillutla, M.M., Murnighan, J.K., 1996. Unfairness, anger, and spite: emotional rejections of ultimatum offers. Organizational Behavior and Human Decision Processes 68 (3), 208–224.

[28] Potegal, M., Stemmler, G., 2010. Cross-disciplinary views of anger: consensus and controversy, in: Potegal, M., Spielberger, C., Stemmler, G. (Eds.), International Handbook of Anger: Constituent and Concomitant Biological, Psychological, and Social Processes, Springer, New York, pp. 3–8.

[29] Rand, D.G., Fudenberg, D., Dreber, A., 2015. It's the thought that counts: the role of intentions in noisy repeated games. Journal of Economic Behavior and Organization 116, 481–499

[30] Ranson, M., 2014. Crime, weather and climate change. Journal of Environmental Economics and Management 67 (3), 274–302.

[31] Robinson, M., Clore, G., 2002. Belief and feeling: evidence for an accessibility model of emotional self-report. Psychological Bulletin 128, 934–960.

[32] Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the ultimatum game. Science 300, 1755–1758.

[33] Smith, C.A., Ellsworth, P.C., 1985. Patterns of cognitive appraisal in emotion. Journal of Personality and Social Psychology 48 (4), 813–838.

[34] Tan, H.-T., Lipe, M.G., 1997. Outcome effects: the impact of decision process and outcome controllability. Journal of Behavioral Decision Making 10 (4), 315–325.

[35] Wranik, T., Scherer, K.R., 2010. Why do I get angry? A componential appraisal approach, in: Potegal, M., Spielberger, C., Stemmler, G. (Eds.), International Handbook of Anger: Constituent and Concomitant Biological, Psychological, and Social Processes, Springer, New York, pp. 243–266.

Appendix A: Additional data

Table 5: All punishment decisions in the experiment (by treatment).

Punishment (player **a**)	*Simple* (Frequency)	*Blame-behavior* (Frequency)	*Simple-strategy* (Frequency)	*Control* (Frequency)
0	22	24	23	20
1	0	1	0	1
2	3	0	2	0
3	0	1	0	1
4	0	1	0	1
5	0	2	0	0
6	0	0	1	0
7	0	0	0	0
8	0	0	0	1
9	0	0	0	0
10	1	0	1	0
Total (N)	26	29	27	24
Average punishment	0.62	0.62	0.74	0.67

Appendix B: Instructions for the *Simple* treatment

General instructions for participant A

Welcome to the experiment and thank you for participating!
Please do not talk to other participants.

You are about to take part in an economics experiment. If you read the following instructions carefully, you can – depending on your decisions – earn money in addition to the SEK 50 you will receive for being part of the experiment. The amount of money that you earn with your decisions will be added up and paid to you in cash at the end of the experiment. These instructions are solely for your private information. You are not allowed to communicate during the experiment. Violation of this rule will lead to exclusion from the experiment and all payments. If you have questions, please raise your hand and we will come to you. We will only answer your questions in private.

We will not speak in terms of Swedish kronor during the experiment, but rather of points. Your income will first be calculated in points. At the end of the experiment, the total amount of points earned will be converted to kronor at the following rate: 1 point = 1 SEK.

We will explain the exact experimental procedure on the next few pages.

The experiment

At the beginning of the experiment, you will be randomly matched with one other participant in the experiment. You will never find out the identity of this person, not even after the experiment. In the same way, the person matched with you will never find out your identity.

There are two types of participants in this experiment: participants A and B. **You are a participant A.** The person matched with you is a participant B. We will refer to you and your participant B as a **group** and each group will be assigned a specific number (a group number).

There are **two possible situations** in the experiment:

- In situation 1, you receive 100 points and participant B receives 100 points.

- In situation 2, you receive 10 points and participant B receives 10 points.

At the beginning of the experiment, the experimenter will randomly assign your group either situation 1 or situation 2 by tossing a 6-sided die.

- If your group number is an **even** number (2, 4, 6, 8 etc.): situation 1 will be assigned if the thrown number is also an even number (2, 4 or 6) and situation 2 will be assigned if the thrown number is an odd number (1, 3 or 5).

- If your group number is an **odd** number (1, 3, 5, 7 etc.): situation 1 will be assigned if the thrown number is also an odd number (1, 3 or 5) and situation 2 will be assigned if the thrown number is an even number (2, 4 or 6).

This means that situation 1 and situation 2 are equally likely to be assigned (each situation will be assigned with a probability of 50%).

Example 1: If your group number is 8 (an even number) and the number on the die is 4 (also an even number), then situation 1 is assigned.

Example 2: If your group number is 4 (an even number) and the number on the die is 3 (an odd number), then situation 2 is assigned.

Example 3: If your group number is 3 (an odd number) and the number on the die is 1 (also an odd number), then situation 1 is assigned.

If situation 1 is assigned, the experiment is over. Your income is 100 points and participant B' income is 100 points.

If situation 2 is assigned, as participant A you have the possibility to deduct up to a total of 10 points from your participant B. Each point deducted will cost you one point. When you have made your choice, the experiment is over.

Example 1: Situation 2 is assigned and you (as participant A) give up 4 points to deduct 4 points from your participant B. The following payment will then result:

- Your points (as participant A): $10 - 4 = 6$.

- Participant B's points: $10 - 4 = 6$.

Example 2: Situation 2 is assigned and you (as participant A) give up 0 points to deduct 0 points from your participant B. The following payment will then result:

- Your points (as participant A): $10 - 0 = 10$.

- Participant B's points: $10 - 0 = 10$.

[*Only in the instructions for participant B:* **In situation 2 you will be asked to estimate the number of points your participant A will deduct from you. You will be paid for the accuracy of your estimate: If your estimate is exactly right, you will get 5 points in addition to your other income from the experiment. If your estimate is not exactly right, you will not get any additional points.**

Example 1: You estimate that your participant A will deduct 4 points from you and this estimate is correct. The following payment will then result:

- Your points (as participant B): $10 - 4 + 5 = 11$.

- Participant A's points: $10 - 4 = 6$.]

Procedures in chronological order

1. The experiment begins. Your group is randomly assigned situation 1 or situation 2. The experimenter will toss the die and one randomly selected participant will verify the outcome. Together they will enter the number on this participant's computer screen.

2a. If situation 1 is assigned, the experiment is over. Your income is 100 points and participant B's income is 100 points. You will receive payment for the points earned plus your show-up fee in cash.

2b. If situation 2 is assigned, as participant A you have the possibility to deduct up to a total of 10 points from your participant B. Each point you deduct will cost you one point. Once you have made your decision, the experiment is over. Your income is 10 points minus the amount (if anything) you deducted from participant B, and participant B's income is 10 points minus the amount (if anything) you deducted from him or her. You will receive payment for the points earned plus your show-up fee in cash.

[*In the instructions for participant B, 2b instead reads:* If situation 2 is assigned, participant A has the possibility to deduct up to a total of 10 points from you. Each point deducted will cost him or her one point. You will estimate the number of points he or she will deduct from you. Once you have made your decision (and your participant A has made his or her decision), the experiment is over. Your income is 10 points minus the amount (if anything) participant A deducted from you plus 5 points if you estimated this amount correctly, and participant A's income is 10 points minus the amount (if anything) he or she deducted from you. You will receive payment for the points earned plus your show-up fee in cash.]

Do you have any questions?

Appendix C: Instructions for the *Blame-behavior* treatment

General instructions for participant A

Welcome to the experiment and thank you for participating!
Please do not talk to other participants.

You are about to take part in an economics experiment. If you read the following instructions carefully, you can – depending on your decisions – earn money in addition to the SEK 50 you will receive for being part of the experiment. The amount of money that you earn with your decisions will be added up and paid to you in cash at the end

of the experiment. These instructions are solely for your private information. You are not allowed to communicate during the experiment. Violation of this rule will lead to exclusion from the experiment and all payments. If you have questions, please raise your hand and we will come to you. We will only answer your questions in private.

We will not speak in terms of Swedish kronor during the experiment, but rather of points. Your income will first be calculated in points. At the end of the experiment, the total amount of points earned will be converted to kronor at the following rate: 1 point = 1 SEK.

We will explain the exact experimental procedure on the next few pages.

The experiment

At the beginning of the experiment, you will be randomly matched with one other participant in the experiment. You will never find out the identity of this person, not even after the experiment. In the same way, the person matched with you will never find out your identity.

There are two types of participants in this experiment: participants A and B. **You are a participant A.** The person matched with you is a participant B. We will refer to you and your participant B as a **group** and each group will be assigned a specific number (a group number).

There are **two possible situations** in the experiment:

- In situation 1, you receive <u>100 points</u> and participant B receives <u>100 points.</u>

- In situation 2, you receive <u>10 points</u> and participant B receives <u>10 points.</u>

At the beginning of the experiment, the experimenter will randomly assign your group either situation 1 or situation 2 by tossing a 6-sided die. Before the experimenter tosses the die, your participant B will **guess** whether the thrown number will be **even** (2, 4 or 6) or **odd** (1, 3 or 5).

- If the guess is **correct,** situation 1 will be assigned.

- If the guess is **wrong,** situation 2 will be assigned.

This means that situation 1 and situation 2 are equally likely to be assigned (each situation will be assigned with a probability of 50%).

Example 1: If your participant B guessed that the thrown number would be even, and the number on the die is 4 (also an even number so the guess was <u>correct</u>), then situation 1 is assigned.

Example 2: If your participant B guessed that the thrown number would be even, and the number on the die is 3 (an odd number so the guess was <u>wrong</u>), then situation 2 is assigned.

Example 3: If your participant B guessed that the thrown number would be odd, and the number on the die is 1 (also an odd number so the guess was <u>right</u>), then situation 1 is assigned.

If situation 1 is assigned, the experiment is over. Your income is 100 points and participant B' income is 100 points.

If situation 2 is assigned, as participant A you have the possibility to deduct up to a total of 10 points from your participant B. Each point deducted will cost you one point. When you have made your choice, the experiment is over.

Example 1: Situation 2 is assigned and you (as participant A) give up 4 points to deduct 4 points from your participant B. The following payment will then result:

- Your points (as participant A): $10 - 4 = 6$.

- Participant B's points: $10 - 4 = 6$.

Example 2: Situation 2 is assigned and you (as participant A) give up 0 points to deduct 0 points from your participant B. The following payment will then result:

- Your points (as participant A): $10 - 0 = 10$.

- Participant B's points: $10 - 0 = 10$.

[*Only in the instructions for participant B:* **In situation 2 you will be asked to estimate the number of points your participant A will deduct from you. You will be paid for the accuracy of your estimate: If your estimate is exactly right, you will get 5 points in addition to your other income from the experiment.**

If your estimate is not exactly right, you will not get any additional points.

<u>Example 1</u>: You estimate that your participant A will deduct 4 points from you and this estimate is correct. The following payment will then result:

- Your points (as participant B): $10 - 4 + 5 = 11$.

- Participant A's points: $10 - 4 = 6$.]

Procedures in chronological order

1. The experiment begins. Your group is randomly assigned situation 1 or situation 2. Your participant B will guess whether the thrown number will be even or odd and the guess will be reported to you (on your computer screen). Then the experimenter will toss the die and one randomly selected participant will verify the outcome. Together they will enter the number on this participant's computer screen.

2a. If situation 1 is assigned, the experiment is over. Your income is 100 points and participant B's income is 100 points. You will receive payment for the points earned plus your show-up fee in cash.

2b. If situation 2 is assigned, as participant A you have the possibility to deduct up to a total of 10 points from your participant B. Each point you deduct will cost you one point. Once you have made your decision, the experiment is over. Your income is 10 points minus the amount (if anything) you deducted from participant B, and participant B's income is 10 points minus the amount (if anything) you deducted from him or her. You will receive payment for the points earned plus your show-up fee in cash.

[*In the instructions for participant B, 2b instead reads:* If situation 2 is assigned, participant A has the possibility to deduct up to a total of 10 points from you. Each point deducted will cost him or her one point. You will estimate the number of points he or she will deduct from you. Once you have made your decision (and your participant A has made his or her decision), the experiment is over. Your income is 10 points minus the amount (if anything) participant A deducted from you plus 5 points if you estimated this amount correctly, and participant A's income is 10 points minus the amount (if anything) he or she deducted from you. You will receive payment for the points earned plus your show-up fee in cash.]

Do you have any questions?

Appendix D: Instructions for the *Simple-strategy* treatment

Welcome to the experiment and thank you for participating!
Please do not talk to other participants.

You are about to take part in an economics experiment. If you read the following instructions carefully, you can – depending on your decisions – earn money in addition to the SEK 50 you will receive for being part of the experiment. The amount of money that you earn with your decisions will be added up and paid to you in cash at the end of the experiment. These instructions are solely for your private information. You are not allowed to communicate during the experiment. Violation of this rule will lead to exclusion from the experiment and all payments. If you have questions, please raise your hand and we will come to you. We will only answer your questions in private.

We will not speak in terms of Swedish kronor during the experiment, but rather of points. Your income will first be calculated in points. At the end of the experiment, the total amount of points earned will be converted to kronor at the following rate: 1 point = 1 SEK.

We will explain the exact experimental procedure on the next few pages.

The experiment

At the beginning of the experiment, you will be randomly matched with one other participant in the experiment. You will never find out the identity of this person, not even after the experiment. In the same way, the person matched with you will never find out your identity.

There are two types of participants in this experiment: participants A and B. **You are a participant A.** The person matched with you is a participant B. We will refer to you and your participant B as a **group** and each group will be assigned a specific number (a group number).

There are **two possible situations** in the experiment:

- In situation 1, you receive <u>100 points</u> and participant B receives <u>100 points</u>.

- In situation 2, you receive <u>10 points</u> and participant B receives <u>10 points.</u>

At the beginning of the experiment, the experimenter will randomly assign your group either situation 1 or situation 2 by tossing a 6-sided die.

- If your group number is an **even** number (2, 4, 6, 8 etc.): situation 1 will be assigned if the thrown number is also an even number (2, 4 or 6) and situation 2 will be assigned if the thrown number is an odd number (1, 3 or 5).

- If your group number is an **odd** number (1, 3, 5, 7 etc.): situation 1 will be assigned if the thrown number is also an odd number (1, 3 or 5) and situation 2 will be assigned if the thrown number is an even number (2, 4 or 6).

This means that situation 1 and situation 2 are equally likely to be assigned (each situation will be assigned with a probability of 50%).

<u>Example 1:</u> If your group number is 8 (an even number) and the number on the die is 4 (also an even number), then situation 1 is assigned.

<u>Example 2:</u> If your group number is 4 (an even number) and the number on the die is 3 (an odd number), then situation 2 is assigned.

<u>Example 3:</u> If your group number is 3 (an odd number) and the number on the die is 1 (also an odd number), then situation 1 is assigned.

If situation 1 is assigned, the experiment is over. Your income is 100 points and participant B' income is 100 points.

If situation 2 is assigned, as participant A you have the possibility to deduct up to a total of 10 points from your participant B. Each point deducted will cost you one point. When you have made your choice, the experiment is over.

<u>Example 1:</u> Situation 2 is assigned and you (as participant A) give up 4 points to deduct 4 points from your participant B. The following payment will then result:

- Your points (as participant A): $10 - 4 = 6$.

- Participant B's points: $10 - 4 = 6$.

Example 2: Situation 2 is assigned and you (as participant A) give up 0 points to deduct 0 points from your participant B. The following payment will then result:

- Your points (as participant A): $10 - 0 = 10$.

- Participant B's points: $10 - 0 = 10$.

[*Only in the instructions for participant B:* **In situation 2 you will be asked to estimate the number of points your participant A will deduct from you. You will be paid for the accuracy of your estimate: If your estimate is exactly right, you will get 5 points in addition to your other income from the experiment. If your estimate is not exactly right, you will not get any additional points.**

Example 1: You estimate that your participant A will deduct 4 points from you and this estimate is correct. The following payment will then result:

- Your points (as participant B): $10 - 4 + 5 = 11$.

- Participant A's points: $10 - 4 = 6$.]

We will ask you to make your choice (how many points to deduct if situation 2 is assigned) <u>before</u> **your group is randomly assigned situation 1 or situation 2 (by the experimenter tossing the die). If situation 2 is assigned, your choice will be implemented (and if situation 1 is assigned your income will be 100 points and participant B's income 100 points regardless of the choice you made).**

Example 1: You (as participant A) give up 4 points to deduct 4 points from your participant B in case situation 2 is assigned. It is then randomly determined that situation 1 will apply (your choice is not implemented). The following payment will then result:

- Your points (as participant A): 100.

- Participant B's points: 100.

Example 2: You (as participant A) give up 4 points to deduct 4 points from your participant B in case situation 2 is assigned. It is then randomly determined that situation 2 will apply (your choice is implemented). The following payment will then result:

- Your points (as participant A): $10 - 4 = 6$.

- Participant B's points: $10 - 4 = 6$.

Procedures in chronological order

1. The experiment begins. As participant A you have the possibility to deduct up to a total of 10 points from your participant B if situation 2 is assigned. Each point you deduct will cost you one point. You will make your choice now and it will be implemented if it is randomly determined that situation 2 will apply.

2. Your group is randomly assigned situation 1 or situation 2. The experimenter will toss the die and one randomly selected participant will verify the outcome. Together they will enter the number on this participant's computer screen.

3a. If situation 1 is assigned, your income is 100 points and participant B's income is 100 points. You will receive payment for the points earned plus your show-up fee in cash.

3b. If situation 2 is assigned, your choice will be implemented. Your income is 10 points minus the amount (if anything) you deducted from participant B, and participant B's income is 10 points minus the amount (if anything) you deducted from him or her. You will receive payment for the points earned plus your show-up fee in cash.

[*In the instructions for participant B, 3b instead reads:* If situation 2 is assigned, your choice will be implemented. Your income is 10 points minus the amount (if anything) participant A deducted from you plus 5 points if you estimated this amount correctly, and participant A's income is 10 points minus the amount (if anything) he or she deducted from you. You will receive payment for the points earned plus your show-up fee in cash.]

Do you have any questions?

Appendix E: Instructions for the *Control* treatment

General instructions for participant A

Welcome to the experiment and thank you for participating!
Please do not talk to other participants.

You are about to take part in an economics experiment. If you read the following instructions carefully, you can – depending on your decisions – earn money in addition to the SEK 50 you will receive for being part of the experiment. The amount of money that you earn with your decisions will be added up and paid to you in cash at the end of the experiment. These instructions are solely for your private information. You are not allowed to communicate during the experiment. Violation of this rule will lead to

exclusion from the experiment and all payments. If you have questions, please raise your hand and we will come to you. We will only answer your questions in private.

We will not speak in terms of Swedish kronor during the experiment, but rather of points. Your income will first be calculated in points. At the end of the experiment, the total amount of points earned will be converted to kronor at the following rate: 1 point = 1 SEK.

We will explain the exact experimental procedure on the next few pages.

The experiment

At the beginning of the experiment, you will be randomly matched with one other participant in the experiment. You will never find out the identity of this person, not even after the experiment. In the same way, the person matched with you will never find out your identity.

There are two types of participants in this experiment: participants A and B. **You are a participant A.** The person matched with you is a participant B.

At the beginning of the experiment, you receive <u>10 points</u> and participant B receives <u>10 points</u>.

As participant A you have the possibility to deduct up to a total of 10 points from your participant B. Each point deducted will cost you one point. When you have made your choice, the experiment is over.

<u>Example 1</u>: You (as participant A) give up 4 points to deduct 4 points from your participant B. The following payment will then result:

- Your points (as participant A): $10 - 4 = 6$.

- Participant B's points: $10 - 4 = 6$.

<u>Example 2</u>: You (as participant A) give up 0 points to deduct 0 points from your participant B. The following payment will then result:

- Your points (as participant A): $10 - 0 = 10$.

- Participant B's points: $10 - 0 = 10$.

[*Only in the instructions for participant B:* **You will be asked to estimate the number of points your participant A will deduct from you. You will be paid for the accuracy of your estimate: If your estimate is exactly right, you will get 5 points in addition to your other income from the experiment. If your estimate is not exactly right, you will not get any additional points.**

Example 1: You estimate that your participant A will deduct 4 points from you and this estimate is correct. The following payment will then result:

- Your points (as participant B): $10 - 4 + 5 = 11$.

- Participant A's points: $10 - 4 = 6$.]

Procedures in chronological order

The experiment begins. You receive 10 points and participant B receives 10 points. As participant A you have the possibility to deduct up to a total of 10 points from your participant B. Each point you deduct will cost you one point. Once you have made your decision, the experiment is over. Your income is 10 points minus the amount (if anything) you deducted from participant B, and participant B's income is 10 points minus the amount (if anything) you deducted from him or her. You will receive payment for the points earned plus your show-up fee in cash.

[*In the instructions for participant B, the above paragraph is replaced by the following:* The experiment begins. You receive 10 points and participant A receives 10 points. Participant A has the possibility to deduct up to a total of 10 points from you. Each point deducted will cost him or her one point. You will estimate the number of points he or she will deduct from you. Once you have made your decision (and your participant A has made his or her decision), the experiment is over. Your income is 10 points minus the amount (if anything) participant A deducted from you plus 5 points if you estimated this amount correctly, and participant A's income is 10 points minus the amount (if anything) he or she deducted from you. You will receive payment for the points earned plus your show-up fee in cash.]

Do you have any questions?

Paper IV

Physician Behavior and Conditional Altruism

The effects of payment system and uncertain health benefit

Emil Persson[*]

Department of Economics, University of Gothenburg, Sweden

Abstract

This paper studies physician behavior using laboratory experiments. The focus is on the effect of the payment system (fee-for-service vs. capitation) on physicians' treatment decisions. More specifically, I investigate the altruistic behavior of physicians and whether this behavior is affected by payment system and uncertainty in health outcome. The experiment shows that many physicians are altruistic toward their patients but also that the degree of altruism varies across patients with different medical needs, indicating that physicians condition their altruism on certain patient characteristics. Moreover, patients are overtreated in fee-for-service payment systems to the same extent as they are undertreated in capitation systems, and this result extends into domains of risk and uncertainty in patient health. Interestingly, the type classification of physicians based on conditional altruism is generally unaffected by payment system; the common categorization is that physician altruism is guided by severity of illness, both under capitation and fee-for-service.

---

[*] Department of Economics, University of Gothenburg, Box 640, SE 405 30 Gothenburg, Sweden; Tel.: +46 31 786 2927; emil.persson@economics.gu.se

1. Introduction

Uncertainty in health and altruistic behavior by doctors are two fundamental aspects in the organization of health care (Arrow, 1963). Individuals are uncertain about the incidence of and recovery from diseases and illnesses, and also about the availability of treatments and their effectiveness for health improvements. Naturally, physicians have greater knowledge about the alternative treatments and their consequences for recovery. This information asymmetry between patients and physicians is a traditional principal-agent problem. Thus, the physicians will to a large extent, if not fully, influence the quantity and quality of the health care provided. Their effort is influenced by their preferences for own income as well as altruism[1] toward the patient, which are together affected by the prevailing payment (remuneration) system (see, e.g., McGuire, 2000 for a general overview of payment systems). Due to budget constraints, a central issue in the provision of health care is allocation. Thus, many countries implement prioritization rules based on general ethical principles (e.g., Sabik and Lie, 2008; Shah, 2009), and physicians are incentivized to provide medical treatment that is in each patient's best interest given these rules. As physicians are the gate keepers to the health care system, it is important for policy makers to understand how physicians trade off own income and altruism and how their behavior in this way corresponds to general principles for priority setting when deciding about quality and quantity of health care provided. It is particularly important that the policy makers understand how the physicians' behavior is affected by the payment system and the presence of uncertainty in the effectiveness of medical treatment. The objective of the present paper is to investigate the altruistic behavior of doctors and whether their behavior is affected by the payment system and uncertainty in health outcome. To this end, I conduct an experiment exploring the causal effect of payment system and health outcome, which in this paper is assumed to depend on the quantity of treatment provided by the physician.

The experiment shows that many physicians are altruistic toward their patients but also that the degree of altruism varies across patients with different medical needs: physicians provide more adequate treatment to patients in greater need of medical care, i.e., they provide further away from own payoff maximization in treating patients with more urgent medical conditions (e.g., patients with worse no-treatment profiles). This indicates that physicians condition their altruism on certain characteristics of the

---

[1] Physician altruism relates to the broader concept of intrinsic motivation. Intrinsic motivation generally refers to actions from which individuals derive internal rewards, whereas extrinsic motivation concerns behavior motivated by external rewards, such as monetary compensation (e.g., Deci, 1971; Deci and Ryan, 1985; Frey and Oberholzer-Gee, 2001; Bénabou and Tirole, 2003).

patients. Interestingly, the type classification of physicians based on such conditional altruism is generally unaffected by payment system; the common categorization is that physician altruism is guided by severity of illness, both under capitation and fee-for-service. The results also show a strong effect of payment system on physicians' provision behavior in general. Consistent with previous studies, patients are undertreated in capitation payment systems and overtreated in fee-for-service systems. With the present study, this result is extended into the domains of risk and uncertainty in the outcome of medical treatment for the patient. While the overall pattern of provision behavior is consistent across domains, detailed analysis using subjects' generic risk and uncertainty preferences reveals individual heterogeneity in their reactions to the introduction of risk and uncertainty in patient health.

The rest of the paper is organized as follows: Section 2 discusses the related literature, Section 3 describes the experimental design, Section 4 provides the experimental results, and Section 5 concludes the paper.

2. Background

Medical costs are increasing substantially, and this is to a large extent driven by an ageing population and rapid technological growth in the health care sector (e.g., Newhouse, 1992; Cutler, 2002). This poses a challenge for health care systems around the world and raises important questions regarding how to organize the health care system, with efficiency and priority setting often being the most important issues to consider. Monetary incentives have become an important component in the design of the health care system, since policy makers must rely on the agency of physicians in implementing their goals and preferred outcomes in health care. A central element in the quality of medical care is that physicians provide medical treatment that is in patients' best interest, given existing rules and regulations for the health care sector. In other words, the health system depends on altruistic physicians who provide good and appropriate medical care for the patients. An important question is therefore how to design physician payment systems in order to achieve the best possible outcome along this dimension.

Despite the increased reliance on monetary incentives in the form of different types of payment systems for providers of health care, relatively little is known empirically about the effects of these systems on physicians' decision making at patient level. A key reason for this is that it is notoriously difficult to find exogenous variation in physician payment system using naturally occurring data. In other words, I would like to investigate the causal effects and not only correlation between two factors, for example

between payment system and quantity of treatment for patients.[2] This is also the most relevant type of information for policy makers. However, it is very challenging to conduct field experiments, not least from an ethical perspective since the physical health of the participants would be affected (without their consent), and only a few studies exist (e.g., Krasnik et al., 1990; Gaynor and Gertler, 1995; Devlin and Sarma, 2008; Shigeoka and Fushimi, 2014). An emerging literature in behavioral and experimental health economics has used economic experiments in order to circumvent these problems of identification. Hennig-Schmidt et al. (2011) introduced a novel experimental framework facilitating detailed investigation of the causal impact of payment system on physician behavior. In their experiment, subjects take on the role of physicians and decide on the provision of medical care for different types of patients who are identical in all respects other than the degree to which a given level of medical treatment affects their health. One of the key benefits of using a lab experiment is that there exists a theoretically known and in my case unique optimal quantity of medical care that maximizes the health benefit in each patient.

One common way to model the physician's behavior is based on Ellis and McGuire (1986). In their model, the utility function of a physician consists of two components, own income and health outcome for the patient, which capture the physician's preference for own income and altruism toward the patient, respectively. In a capitation remuneration system, the physician receives a fixed payment related to the people she is assumed to be responsible to care for, and then makes an effort to improve their health. Thus, in the capitation system the emphasis is on the patient's health and the physician's altruism as expressed by giving up own income to improve the health of the patient. In contrast, in a fee-for-service remuneration system, physicians are paid for each amount of health care provided. As can easily be seen, these two systems consider the physician's preference for own income differently. In the extreme case of a physician motivated exclusively by own income maximization, the two systems provide completely different amounts of treatment. Under capitation, there will not be any treatment and hence the patient will be undertreated, while under fee-for-service the patient will receive the maximum treatment, or until the point where the marginal utility of effort equals the utility of income provided by that unit of effort for the physician. In a fee-for-service system, this could result in treatment levels that might be both harmful and a waste of resources, which relates to what has been coined supplier-induced demand (e.g., Fuchs, 1978; Dranove and Wehner, 1994; Gruber and Owings, 1996;

---

[2] For a general discussion on causality and experimental and survey methods, see Falk and Heckman (2009).

Shigeoka and Fushimi, 2014). In the other extreme case, where the physician is only affected by altruism,[3] both payment systems result in optimal provision of health care for the patient. More likely, however, is that we have intermediate cases where the physician derives utility both from own income and altruism, in line with the discussion already in Arrow (1963). The relative degree of physicians' preference for own income in relation to their altruism influences which of the two remuneration systems is preferred from a societal perspective.

A robust finding in the experimental literature on physician behavior is that physicians provide significantly more medical care under fee-for-service payment systems than under capitation payment systems, to the extent that patients are overtreated under fee-for-service and undertreated under capitation (Hennig-Schmidt et al., 2011; Brosig-Koch et al., 2013; Keser et al, 2014; Hennig-Schmidt and Wiesen, 2014; Brosig-Koch et al., 2015a,b). Among these studies, the experimental designs in Brosig-Koch et al. (2013) and Brosig-Koch et al. (2015a,b) facilitate analysis of whether physicians provide better medical treatment under one payment system or the other, and capitation seems to be the marginally better system in this respect. In a re-analysis of the data from Hennig-Schmidt et al. (2011), Godager and Wiesen (2013) focus on estimating the relative weight attached to own income and altruism. Their results confirm that most subjects assign a weight to both components, but there is large heterogeneity in the weights assigned.

An interesting possibility concerning physician altruism is that it might vary across different types of patients. For example, physicians might be more altruistic toward patients with more severe illnesses, or they might be more altruistic toward patients with a high capacity to benefit from treatment. These are two different principles for allocation of health care, and they can be traced back to priority-setting principles based on medical need: The capacity-to-benefit principle implies a goal to maximize the overall level of health in the population (e.g., maximize the total number of quality-adjusted life years, QALYs), whereas the severity-of-illness principle implies that health maximization should be weighted by medical condition (e.g., Dolan and Olsen, 2001; Dolan et al., 2005). In practice, capacity to benefit has been the overarching goal in for example the UK, using QALY maximization as an explicit criterion, whereas severity of illness has been important in Norwegian, Swedish, and German health policy (Shah, 2009).

---

[3] One example of this is Doctors Without Borders.

5

How physicians' behavior corresponds to different principles for priority setting therefore has implications for how well the quality and quantity of health care provided correspond with existing rules and guidelines established for the health care sector. Different principles for allocation according to medical need yield different behavior, which can be traced back to different rules for priority setting (Culyer and Wagstaff, 1993; Olsen, 1997; Cookson and Dolan, 2000). Based on the taxonomy in Williams and Cookson (2000), we can identify three general principles for provision according to medical need (two of which have already been mentioned). First, a principle to prioritize based on *severity of illness* gives precedence to the patient with the worst no-treatment profile, i.e., the one with the worst outlook in the absence of treatment. This principle has been justified on grounds of equality (e.g., Williams, 1962), and, for urgent health problems (such as risk of death), based on a "rule of rescue" (Cookson and Dolan, 2000). Second, a *capacity-to-benefit* principle places the greatest weight on patients who will benefit the most from treatment and it is thus justified on grounds of maximizing the overall level of population health. Finally, one could provide health care with an aim to *equalize ex-post health*, a principle that can be justified by for example the "fair innings" argument, which stresses everyone's right to a similarly normal span of life years in good health (Williams, 1997).

For a systematic and detailed investigation of the influence of patients' medical need on physician decision making, an experimental approach is suitable since it facilitates controlled ceteris paribus variation in patient characteristics and medical conditions, something that would be extremely difficult to accomplish using naturally occurring data. Based on an extension of the design in Hennig-Schmidt et al. (2011), the present paper introduces a type classification of physicians in terms of conditional altruism, capturing how well their treatment decisions in the experiment align with the different principles for priority setting discussed above: severity of illness, capacity to benefit, and equalization of ex-post health (Williams and Cookson, 2000).

There is often uncertainty about patients' recovery from illness as well as the effectiveness of medical treatment, and this could affect physician behavior and altruism toward the patient. For example, in the event of uncertainty, some physicians might focus on the worst possible outcomes and thus become more cautious in their decisions compared with when there is little or no uncertainty regarding the effect of treatment for the patient. Generally speaking, assuming risk aversion we expect physicians in the experiment to on average provide medical treatment closer to patients' optimal treatment levels when the treatment outcome is risky or uncertain than when it is certain. Since previous experimental research based on the framework introduced by Hennig-Schmidt et al. (2011) has focused exclusively on situations when the outcome

of medical treatment is known for sure (it is deterministic), introducing the realistic aspects of risk and uncertainty in patient health seems very important for the analysis of physician response to payment system.

In summary, I implement two novel extensions to the experimental framework introduced by Hennig-Schmidt et al. (2011) and extended by Brosig-Koch et al. (2015c). This contributes to fill important gaps in the literature seeking to understand the impact of payment systems and the effect of risk and uncertainty in health benefits of treatments on physician behavior.

3. Experiment

3.1 Design

The focus of the paper is on the effect of payment system (fee-for-service vs. capitation) on physicians' treatment decisions. More specifically, I investigate the altruistic behavior of physicians and whether their behavior is affected by payment system, uncertainty in health outcome, and differences in patients' medical needs. In the experiment, physicians are paid either through *capitation* or *fee-for-service*, and they meet patients who differ systematically from each other (i) in how certain their health benefits from treatment are (*deterministic*, *risky*, or *uncertain*) and (ii) in their *medical needs*.

The experimental design builds on the framework introduced by Hennig-Schmidt et al. (2011) and extended by Brosig-Koch, Hennig-Schmidt, Kairies-Schwarz, and Wiesen in a series of recent papers, e.g., Brosig-Koch et al. (2015c). I stay as close as possible to the latter version both in terms of experimental instructions and choice of parameters. In their (and my) experiment, subjects in the role of physicians decide on the quantity of medical treatment $q \in \{0, 1, \ldots, 10\}$ for different types of patients.[4] In treating the patients, physicians incur costs following a convex cost function and they also receive a payment as either a fixed sum up front (*capitation*) or a sum that varies with the level of medical treatment they provide (*fee-for-service*).

In the experiment, each physician met patients with five different health profiles (A, B, C, D, and E). Each patient in the experiment is characterized by a health benefit

---

[4] In reality, the treatment decision is ultimately made by the patient after trusting that the physician has provided accurate information about treatment possibilities and outcomes. However, since the information given is based on the physician's assessments and decisions, in the experiment I assume that the physician is the actual decision maker.

function mapping the amount of medical treatment (q) into health benefit expressed in monetary terms. Figure 1 displays the benefit functions for the health profiles used in the experiment. For instance, if the physician gives a patient with health profile A five units of medical treatment (q=5), the resulting patient health benefit is 8 taler (the experimental currency used, later translated into euros), and the highest health benefit is received if q=7. The other health profiles are interpreted in a similar way.

Figure 1 about here

For each patient there is a unique interior level of medical treatment where the health benefit is maximized, and in either payment system the physician is faced with a tradeoff between patient health and own payoff maximization. Parameters for physicians' cost of treatment and the payment they receive are set such that a purely selfish physician would choose q=0 (provide no medical treatment) under capitation. If the physician is not purely selfish but instead to some degree motivated by altruism toward the patient, he or she would choose q>0, and if sufficiently altruistic the physician would provide the level of medical treatment where the patient's health benefit is maximized. For health profile A, this entails q=7. Under fee-for-service, a purely selfish physician would choose q=10 (i.e., would provide maximal medical treatment in order to maximize her own monetary payoff) whereas an altruistic physician would choose q<10 and if sufficiently altruistic would provide the medical treatment that is optimal for the patient. In order to facilitate comparison of the two different payment systems, the absolute values of marginal profits as well as the maximum profits are equivalent under both payment systems.[5] The benefit functions are symmetric as well, such that a deviation from optimal treatment is equally bad in terms of health benefits whether caused by, say, a two-unit overtreatment or a two-unit undertreatment; in other words, the absolute deviation from optimal treatment is what influences the realized health benefits from the treatments. This is a simplification of reality but allows for clean comparison of different payment systems since the marginal effects (for the patient) of undertreatment and overtreatment are identical (Brosig-Koch et al., 2015a).

There are no real patients present in the lab, but the amounts resulting from physicians' decisions are transferred to a charity that funds certain medical treatments (described further below). These design features were key parts in the novel design introduced by

---

[5] Experiment parameters: The cost function is $c(q) = q^2/10$ and it is the same for all patients; the capitation payment is $R^{CAP} = 10$ for each patient; the fee-for-service payment is $R^{FFS} = 2q$ for each patient.

Hennig-Schmidt et al. (2011) and have since then been adopted by subsequent studies that build on the Hennig-Schmidt et al. design. The donation is crucial in creating a frame where decisions in the lab are made in a medical context. It means that physicians' decisions in the experiment have consequences for real patients outside the lab and thereby they are explicitly incentivized to carefully consider the consequences for each (abstract) patient in the experiment when making their treatment decisions. The charity used was Christoffel Blindenmission Deutschland e.V., an organization that funds treatment of patients with eye cataracts, and the same organization was used as a recipient in previous studies using this design. In order to ensure participants that the donations would in fact be made after the experiment, the experiment instructions informed them that receipts of each transfer would be published on the experimental laboratory's blackboard on a given date (see instructions in the Appendix). This allowed them to verify that the transferred amount corresponded to the specific amount (in euro) resulting from the payoff-relevant decisions made by all participants. This is standard procedure at MELESSA at the University of Munich, where all sessions of the experiment were conducted.

In order to meet my research goals, I extend the design in Brosig-Koch et al. (2015c) in two dimensions. First, I allow for risk and uncertainty in the outcome of medical treatment. This is an extension compared with Brosig-Koch et al. and other previous studies, which only considered deterministic outcomes. Risk relates to situations with objectively known probabilities of outcomes whereas uncertainty concerns situations with vague or unknown probabilities. It is important to consider both risk and uncertainty, since previous research has shown that individuals make different choices under risk than under uncertainty, and that they tend to prefer situations characterized by risk rather than uncertainty even when one takes their subjective beliefs about probabilities into account (Ellsberg, 1961; for a review see, e.g., Camerer and Weber, 1992). Uncertainty is especially relevant in the context of health care since in many situations involving clinical decision making, it is difficult to pin down the exact probability of a successful treatment outcome due to for example the complexity of many illnesses and expert disagreement regarding the effect of many treatments (for a discussion, see, e.g., Berger et al., 2013).

I introduced risk in the deterministic design as follows: For each deterministic health benefit function (corresponding to each of the five health profiles mentioned earlier), a given level of medical treatment (q) yields a *good* outcome for the patient in terms of health benefit with a 0.5 probability and a *bad* outcome for the patient with a 0.5 probability. The good outcome was always one taler above the corresponding outcome in the deterministic case, and, conversely, the bad outcome was always one taler below

the corresponding deterministic outcome. For example, in terms of the deterministic benefit function characterizing the patient with health profile A (see Figure 1), the good outcome for the patient treated with q=5 units of medical care was 9 taler and the bad outcome 7 taler.[6] Uncertainty was introduced in a similar manner, with a good outcome at one taler above the corresponding deterministic outcome and a bad outcome at one taler below; yet in contrast to risk, the probabilities of the outcomes were unknown. Throughout the experiment, risk was represented by an opaque bag consisting of five black and five white balls, and uncertainty was represented by an opaque bag consisting of ten black and white balls, and outcomes were based on betting correctly on one of the two colors. This is similar for instance to the procedure in Sutter et al. (2013).

I also extend the design in Brosig-Koch et al. (2015c) by investigating physician behavior and altruism toward the patients based on a systematic pairwise comparison of different types of benefit functions. I employ a slightly different set of benefit functions compared with Brosig-Koch et al. In my experiment, physicians met patients with five different health profiles (A, B, C, D, and E) which were separated by the shape of the health benefit function. In more detail, the health profiles differ with respect to: (i) the severity of patients' health problems (the intercept of the benefit function), (ii) patients' capacity to benefit from medical treatment (the slope), and (iii) the optimal level of medical treatment (where the health benefit is maximized). Figure 1 shows the benefit functions. In the spirit of Brosig-Koch et al. (2015c), the benefit functions are carefully designed so as to facilitate comparison across the two different payment systems used in the experiment (capitation and fee-for-service). For example, Figure 1 shows that the benefit function for health profile A is the mirror image of the one for profile B. This means that the marginal effect of undertreating A-patients, which is more likely under capitation, is identical to the marginal effect of overtreating B-patients, which is more likely under fee-for-service. A similar relationship between under- and overtreatment exists for benefit functions C, D, and E, since all of them are symmetric around q=5, which is the point where they reach their maximum value.

The analysis of how physician behavior and altruism vary across patients with different

---

[6] This implementation of risk in patient health resembles a situation where the variation in health outcome for the patient can differ across conditions and interventions even when the expected health gain for a single patient is quite similar. In the experiment, the outcome for the patient in the deterministic treatment is known for sure but in the risk treatment there is some variation in that the outcome is either better or worse compared with the deterministic treatment. This is realistic since there is often a varying degree of knowledge about various conditions, and some medical treatments have been more tested than others. The choice of probability at exactly 50-50 is made to facilitate understanding. Other combinations would be interesting as well and also plausible from a real-world perspective, such as a 0.9 probability of a modest gain (relative to the deterministic treatment) and a 0.1 probability of a very serious outcome.

medical needs is based on a comparison of patients with health profiles C and E, which are similar in that their health benefit is maximized at the same level of medical treatment (five units) and that their functions have the same intercept, but different in that the slope of the benefit function is steeper for E than for C. This allows for the following classification of physicians in terms of how well their treatment decisions in the experiment align with the different principles for priority setting according to medical need: First, physicians' behavior corresponds to the *equality principle* if they provide better (i.e., closer to optimal) medical care to C-patients. These patients' benefit function is flatter compared with E patients, which means that their capacity to benefit from medical care is lower and, thus, better medical care should be provided to C if the goal is to equalize health benefit ex post. For instance, four units to C and two units to E would achieve this goal. Second, physicians' behavior corresponds to the principle of *capacity to benefit* if they provide worse medical care to C, since according to this principle medical care should be allocated in order to produce maximal health benefit. When this is the goal, C is at a disadvantage since the marginal increase in health benefit is lower for C than E due to the flatter slope of the benefit function. Third, physicians' behavior corresponds to the *severity-of-illness principle* if they provide equal amounts of medical care to C and E, since both patients' benefit functions have the same intercept and therefore the same no-treatment profile, i.e., their health benefit is identical without medical treatment. In addition, physicians who are purely self-interested would maximize their own income and thus provide maximal medical treatment to both C and E in a fee-for-service payment system but not provide any medical treatment to either C or E in a capitation payment system. Physicians who are sufficiently altruistic toward their patients would give both C and E the optimal medical treatment, i.e., five units.

Two types of sessions were conducted. The only difference between them is whether a *capitation* payment system (CAP sessions) or a *fee-for-service* payment system (FFS sessions) was used. In each session, physicians first made five treatment decisions in the case of deterministic patient health and each patient belonged to a different health profile A–E; then followed another five treatment decisions in the case of risky patient health; and the last five treatment decisions concerned patients with uncertain outcomes of medical treatment. This first part of the experiment is summarized in Table 1.

Table 1 about here

In the second part of the experiment, subjects' individual risk and uncertainty preferences were elicited in a generic setting. I followed Sutter et al. (2013) and used incentivized choice lists where subjects decided between a fixed prospect and a sure

amount of money. The fixed prospect consists of a 50-50 lottery where the outcome is either 0 or 5 taler. The subject is asked to make repeated choices between a sure amount and the fixed prospect. The sure amount in the first choice is 0.25 taler and it is then increased by increments of 0.25 taler until it reaches 5 taler. For example, a risk-neutral subject would choose the fixed prospect when the sure amount is below 2.50 taler, which corresponds to the expected value of the lottery, and the sure amount when this amount is 2.50 taler or higher. Similarly, a risk-averse subject would choose the fixed prospect fewer times than a risk-neutral subject, while the opposite is true for a risk-loving individual. Subjects completed two lists in total: the first concerning a risky prospect (known probabilities) and the second an uncertain prospect (unknown probabilities). Risk was represented by an opaque bag consisting of five black and five white balls and uncertainty was represented by an opaque bag consisting of ten black and white balls, exactly as in part one of the experiment. The money won added to the subjects' total earnings in the experiment. In order to facilitate a gradual transition between physician decision making (first part of the experiment) and elicitation of individual risk and uncertainty preferences in the generic domain (second part of the experiment), subjects answered a few general questions related to health care and medical decision making upon completion of the first part of the experiment.[7]

3.2 Procedural details

The experiment was conducted at MELESSA, University of Munich. It was computerized using z-Tree (Fischbacher, 2007) and participants were recruited from the existing student subject pool using the ORSEE online recruitment system (Greiner, 2015). When subjects arrived, they were welcomed to the lab and randomized into cubicle workstations. They were told that the experiment consisted of several unrelated parts and received written instructions prior to each part (see Appendix). They were given plenty of time to read the instructions and ask questions (in private) before each part began. Before the first part of the experiment, they had to answer several control questions, and the experiment did not start until everyone had answered all questions correctly.

There was no feedback between the different parts of the experiment. At the end of the experiment, subjects were paid for one randomly drawn decision in part one of the experiment, which concerned physician decision making, and for one randomly drawn

---

[7] Additionally, there was a third part with another type of risk experiment. This part is not used in the current study. It was conducted after completion of the other two parts of the experiment and therefore does not affect the results in this paper.

decision in part two of the experiment, which concerned subjects' individual risk and uncertainty preferences in a generic setting. Throughout the experiment, risk was represented by an opaque bag consisting of five black and five white balls and uncertainty was represented by an opaque bag consisting of ten black and white balls. There was one bag for risk and another bag for uncertainty for each of the different parts of the experiment. At the end of the experiment, one randomly selected individual drew a ball from each bag (and another individual announced the outcome). Before the drawing, all subjects had to guess (in private) the color of the ball drawn from each bag and a correct guess resulted in a good outcome and an incorrect guess resulted in a bad outcome for the corresponding decision of that particular individual. In order to ensure a credible process, subjects were told that they could verify the results from the public drawing (and check the composition of each of the different bags) at the end of the experiment.

An experimental session lasted around 75 minutes and a total of 130 students participated in the experiment, 64 in FFS sessions and 66 in CAP sessions.

4. Results

I begin with a description of the average quantities of medical treatment chosen for each patient under FFS and CAP, respectively. These are displayed in Figure 2, together with the optimal quantity of medical treatment required by each patient (dotted line). The figure provides a summary of the experimental results, and two main findings stand out. First, there is a strong effect of payment system on the quantity of medical treatment provided by physicians in the experiment. For example, patient 1 is on average provided 8.03 units of medical treatment in FFS and 2.73 units of medical treatment in CAP. Averaged over all patients, 66% less medical treatment was provided in CAP compared with FFS, and we can see in the figure that patients are on average undertreated in CAP and overtreated in FFS. These results are in line with previous experimental results (e.g., Hennig-Schmidt et al., 2011; Brosig-Koch et al., 2015c). Second, this pattern prevails under risk and uncertainty in the effect of medical treatment on patient health. For example, patients with health profile A (patients 1, 6, and 11) are provided close to 8 units of medical care by FFS physicians in all three health benefit categories, i.e., deterministic, risky and uncertain health benefits. On average, physicians seem to neither increase nor decrease their level of medical treatment compared with the baseline case of treating the same types of patients with deterministic benefit functions.

Figure 2 about here

13

Table 2 shows in detail the effect of payment system on the average deviation from optimal medical treatment. Patients are overtreated by an average of 2.48 units of medical treatment in FFS and undertreated by an average of 2.46 units in CAP. The difference between FFS and CAP is significant at the 1% level (Mann-Whitney U test; $p < 0.01$, $N = 130$). Since the benefit functions are symmetric such that the marginal effects (for the patient) of undertreatment and overtreatment are identical, the absolute deviation from optimal treatment is what influences the realized health benefit for the patient. We can see that the average absolute deviation from optimal treatment is not significantly different between CAP and FFS (Mann-Whitney U test; $p = 0.82$, $N = 130$). The final two columns of Table 2 show the proportion of purely selfish decisions (q=0 in CAP and q=10 in FFS) and the proportion of decisions that were optimal for the patient, respectively. Interestingly, there seem to be more selfish physicians but also more physicians who are purely altruistic in FFS than in CAP, but neither difference is significant in this case. Overall, patients are treated equally well (or badly) in either payment system; i.e., the problems of undertreatment (in CAP) and overtreatment (in FFS) are approximately similar in terms of magnitude, resulting in the same health levels for the patients.[8]

Table 2 about here

Table 3 shows the effect of for sure, risky, and uncertain outcome in health on the choices of quantity of medical treatment measured as the absolute deviation from optimal treatment of patients.[9] In general, there is a small increase in the absolute deviation from optimal treatment when the outcome in patient health is risky or uncertain instead of sure. Since it is a within-subject comparison, I use a signed rank test. The comparisons with the baseline, i.e., deterministic health benefit, are generally insignificant apart from the case of CAP, where there is a difference between uncertain and deterministic health benefits. By and large, the introduction of risk and uncertainty in the effect of medical treatment on patient health does not seem to affect physicians' provision behavior in the experiment.

Table 3 about here

Tables 4 and 5 offer a more detailed analysis of treatment decisions. The tables present the share of subjects providing optimal medical treatment, for each health profile in the

---

[8] The results are similar when disaggregated by health profile (A and B vs. C, D, and E), and this can be seen in the Appendix, Table A1.

[9] See Tables A2 and A3 in the Appendix for a disaggregation by health profile, which yields similar results.

case of deterministic, risky, and uncertain health benefit functions, respectively. To begin with, we can see a difference in the first two health profiles, A and B. This is expected since they are mirror images of each other, with A requiring a larger quantity of medical treatment (for optimal health benefit) than B, thereby making it more costly in terms of foregone income to treat A than B in CAP but less costly in FFS. For example, in deterministic FFS, 43.8% of subjects provide optimal treatment to the patient with health profile A but only 14.1% of subjects do so for health profile B; by contrast, only 4.5% of subjects treat A optimally in CAP but 50% of subjects do so for B. Health profiles C–E are similar to each other in that the patient health benefit is maximized at five units of medical care for all three profiles, and the differences across FFS and CAP are less pronounced as expected since it is equally costly for physicians to provide optimal medical treatment under either payment system. Turning to the effect of risk and uncertainty, we can see that, in FFS, there is a significant difference at the 10% level using Wilcoxon signed-rank tests only for health profiles B and D and only when comparing deterministic and risky health benefit functions. In CAP, there is a significant difference at the 10% level using Wilcoxon signed-rank tests for health profiles B and D when comparing deterministic and uncertain health benefit functions, and there is also a significant difference at the 5% level using a Wilcoxon signed-rank test between deterministic and risky health benefit functions for health profile D. Overall, the share of subjects who provide optimal medical care seems little affected by risk and uncertainty in patient health under both payment systems.

Tables 4 and 5 about here

The main results regarding the effects of payment system and of risk and uncertainty in the effect of medical treatment on patient health are formulated as follows.

Result 1. *There is a strong effect of payment system on physician provision behavior. Patients are undertreated under capitation and overtreated under fee-for-service. Moreover, they are undertreated under capitation to approximately the same degree as they are overtreated under fee-for-service. Patients are thus on average equally well off under both systems.*

Result 2. *The overall pattern of provision behavior is unaffected by the introduction of risk and uncertainty in the effect of medical treatment on patient health.*

The effects of risk and uncertainty are investigated in more detail by using subjects' generic risk and uncertainty preferences, which were elicited in a later (second) part of the experiment. Table 6 compares deterministic and risky health benefits: medical treatment measured as the absolute deviation from optimal treatment is regressed on a

dummy for the five patients with risky benefit functions (*Risk*), on the measure of individual risk preference (*Risk for self*), and on the interaction between these two variables and several control variables (including age, gender, and health status). Standard errors in the regressions are clustered on individuals. *Risk for self* was elicited using incentivized choice lists where subjects decided between a fixed prospect and a sure amount of money that increased along the list. In FFS (columns 1 and 2), there is a small effect (significant at the 10% level) of the introduction of risk, such that physicians on average increase their overtreatment by 0.105 units of medical care; yet there is no effect of subjects' risk preferences. In contrast, in CAP, subjects' risk preferences are correlated with both how well they treat their patients on average (column 3) and how they change their behavior when patient health is risky rather than deterministic (column 4). As indicated by the positive and significant interaction term, introducing risk in the effect of medical treatment on patient health makes subjects who are more risk averse provide medical treatment closer to patients' optimal treatment levels.[10] This effect is partly explained by the fact that these subjects seem to provide worse medical care in the first place, i.e., when patient health is deterministic.

Table 6 about here

Table 7 shows the results from the corresponding analysis of uncertainty in the effect of medical treatment on patient health.[11] The uncertainty treatment is identical to the risk treatment except that the probabilities of the outcome of medical treatment for the patients were unknown. There is an overall and marginally significant effect of uncertainty on physicians' provision behavior in FFS (columns 1 and 2), similar to the effect of risk investigated in Table 6. In CAP, the effect is correlated with individuals' generic uncertainty preferences (column 4): More uncertainty-averse subjects become better doctors and this effect is similar to the one observed when patient health is risky. However, unlike their risk preferences, subjects' uncertainty preferences are not

---

[10] For instance, when the expression in column 4 of Table 6 is evaluated at *Risk for self* = 6, the marginal effect of *Risk* on the absolute deviation from optimal treatment is -0.22 ($p = 0.02$). This would implies that physicians on average provide 0.22 units of medical care closer to patients' optimum when patient health is risky. At *Risk for self* = 4, the marginal effect is -0.38 ($p = 0.02$). (*Risk for self* = 6 means that that subjects switch to the certain amount of money at row 6 in the individual choice task, consistent with risk aversion. In CAP, 20% of the subjects switch here or earlier. Risk-neutral individuals would switch at 10.)

[11] The results for both risk and uncertainty are similar when disaggregated by health profile (A and B vs. C, D, and E) and this can be seen in the Appendix, Tables A4 and A5. Using the level of medical treatment (q) as a dependent variable instead of the absolute deviation from optimal treatment (as in Tables 6 and 7) produces similar results, except that the correlation between the effect of risk (uncertainty) and individuals' generic risk (uncertainty) preferences is somewhat weaker. The effect is similar when using Tobit instead of OLS; these regressions are available upon request.

correlated with the level of medical treatment they provide when patient health is deterministic (column 3).

Table 7 about here

Result 3. *Physicians' generic risk and uncertainty preferences affect their reactions to risk and uncertainty in the effect of medical treatment on patient health under CAP. In particular, more risk-averse subjects provide treatment closer to patients' optimal treatment levels following the introduction of risk in patient health; and more uncertainty-averse individuals provide treatment closer to patient optimum following the introduction of uncertainty in patient health.*

Next, we turn to the analysis of altruism and conditional altruism of physicians. Physician behavior is assumed to be influenced by two key components: their altruism toward the patient and preference for own income. Physicians who are purely selfish choose q=0 and q=10 under CAP and FFS, respectively. Table 8 displays the proportion of subjects who made a selfish decision at least once and also the proportion who were selfish in all their treatment decisions. We restrict attention to the case of deterministic benefit functions. For example, in FFS, 20.3% of the physicians always made a purely selfish decision and 31.3% did so at least once. By comparison, in CAP, 10.6% of the physicians were always selfish and 27.3% were selfish at least once. It is interesting that physicians who were always selfish when deciding for patients with profiles C, D, or E were also selfish in their decisions for patients with profiles A and B, even though it is comparatively less costly to provide one of these patients with optimal medical treatment.[12] The table also shows a similar analysis for the proportion of physicians who made optimal decisions for their patients. For example, 46.9% in FFS and 53.0% in CAP decided optimally at least once, whereas 10.9% in FFS and 3.0% in CAP provided optimal medical treatment for all of their patients. This indicates that many physicians are altruistic but also that the degree of altruism varies across different types of patients. This is further investigated in Table 9, where physicians' provision of medical care to patients from health profiles C and D are compared. The comparison is interesting since based on the discussion about different principles of medical need (equality, capacity to benefit, and severity of illness), we know that C patients are in equal or greater need than D patients: While their capacity to benefit from medical care is identical, they are worse off without treatment (severity-of-illness principle) and also require better

---

[12] Optimal medical treatment is 7 units for patients from A, 3 units for patients from B, and 5 units for patients from C, D, or E. Thus, in CAP it is less costly to provide B with optimal treatment than C, D, or E; and vice versa for FFS where it is less costly to provide A with optimal medical treatment.

treatment in order to equalize health ex post (equality principle). In the table, we can see that patients with health profile C do receive better medical treatment than patients with health profile D. The difference is most pronounced under the fee-for-service payment system, in particular when there is risk or uncertainty in the effect of medical treatment on patient health. For example, with uncertainty in patient health, FFS physicians provided 0.42 units of medical care closer to patient optimum when treating patients with profile C than when treating patients with profile D.

Tables 8 and 9 about here

Result 4. *Many physicians are altruistic toward their patients. The degree of altruism depends on the medical need of the patient: Patients with health profile C are provided better medical care than patients with profile D; the effect is stronger under FFS and increases with risk and uncertainty in the effect of medical treatment on patient health.*

The type classification of physicians based on how their behavior and altruism vary across patients with different medical needs is based on a comparison of medical treatment provided to patients with health profiles C and E. We can identify three broad categories of physicians: (1) purely selfish physicians who always choose so as to maximize own income, (2) purely altruistic physicians who always choose so as to maximize patient health benefit, and (3) physicians who take both of these two perspectives into account, choosing a mix between full altruism and own income maximization. Table 10 shows the distribution of physicians over these three categories based on their decisions in the case of deterministic benefit functions. Interestingly, there are more purely selfish physicians but also more purely altruistic physicians in FFS than in CAP, which thus has a greater share of physicians in the third category, i.e., who seem to be motivated both by own income and altruism. Based on these three categories, the type distribution is significantly different between the two payment systems (Kruskal-Wallis test; $p = 0.05$, $N = 130$).

The only difference between C and E is that the slope of the benefit function is steeper in the latter and this allows for the following more detailed classification of physicians within the mix-category (see Section 3 for details): Physicians' behavior corresponds to the priority principle of *equality ex post* if they provide better (i.e., closer to optimal) medical care to C; physicians' behavior corresponds to the *capacity-to-benefit principle* if they provide worse medical care to C; and, finally, physicians' behavior corresponds to the *severity-of-illness principle* if they provide equal amounts of medical care to C and E. This more detailed information about the distribution of physician types is also displayed in Table 10. Overall, we can see that among those who choose a mix between full altruism and own income maximization, the altruistic part is mostly described by

concerns for severity of illness (38.5% and 47.1% of FFS and CAP subjects, respectively). In general, this principle implies that patients with worse no-treatment profiles should be given priority, even if their capacity to benefit from medical treatment is comparatively low. The capacity-to-benefit principle follows quite closely, with the behavior of 35.9% of FFS subjects and 39.2% of CAP subjects corresponding to this principle. A minority of subjects behave in accordance with the principle of ex-post equality. Moreover, we can see that there are some differences in the distribution of physician types between the two payment systems. In FFS, the distribution is fairly balanced across the different types: 25.6%, 35.9%, and 38.5% of subjects provide treatment corresponding to equality ex-post, capacity to benefit, and severity of illness, respectively. In contrast, in CAP, only a small minority of physicians, i.e., 13.7%, provide treatment corresponding to the equality-ex-post principle, whereas 47.1% make decisions consistent with the severity-of-illness principle. However, these differences in distribution are not statistically significant in a Kruskal-Wallis test, neither when comparing within the full sample ($p = 0.32$, $N = 130$) nor when comparing within the subsample of physicians who belong to the mix-category ($p = 0.23$, $N = 89$). Thus, taken together, physician altruism seems fairly robust to the choice of payment system in the experiment.

Table 10 about here

Result 5. *There are more pure altruists but also more pure selfish physicians in FFS than in CAP, but the more detailed distribution of physician types according to how well their treatment decisions correspond to the different principles for priority setting is not affected by payment system. In both CAP and FFS, altruistic behavior toward the patients is mostly guided by severity of illness.*

5. Discussion and conclusion

As already discussed by Arrow (1963), two key issues for the health care sector are uncertainty in health and altruistic behavior by physicians. In this paper, I use a lab experiment to investigate the altruistic behavior of physicians and whether this behavior is affected by payment system and risk and uncertainty in health outcome. In order to contain increasing medical costs in health care systems around the world, there is a growing focus on using monetary incentives in the provision of health care. The aim is to incentivize physicians to provide medical treatment that is in each patient's best interest, given existing rules and regulations for the health care sector. However, there is limited empirical evidence regarding the direct impact of such incentives on clinical decision making, and thus on the ability of policy makers to use different types of

physician reimbursement schemes in order to implement their goals and preferred outcomes in health care. A key issue is that it is difficult to find exogenous variation in physician payment systems using naturally occurring data, for example in the form of natural experiments or field experiments where the payment systems are varied, which would be needed in order to make causal inference.

This paper uses a lab experiment to circumvent problems of identification and furthermore to provide detailed analysis of individual-level decision making, thereby contributing to the existing literature on effects of payment system on the provision of health care. Building on the experimental framework introduced by Hennig-Schmidt et al. (2011) and used and extended by many researchers, most notably by Brosig-Koch et al. (2015c), I implement two novel extensions, thus investigating (i) how physician altruism is affected by payment system (capitation vs. fee-for-service) and patients' medical needs and (ii) how physicians' response to payment system is affected by risk and uncertainty in the outcome of medical treatment for the patient, which is a fundamental aspect of physicians' decision environment in the real world. This contributes to fill important gaps in the literature seeking to understand the detailed effects of different payment systems on physicians' treatment decisions, where their utility is assumed to consist of two key elements: concern for own income and altruism toward the patient. Altruism among physicians was discussed by Arrow (1963) and then prominently developed by Ellis and McGuire (1986).[13] Godager and Wiesen (2013) analyze physician altruism based on Ellis and McGuire (1986) using the data in Hennig-Schmidt et al. (2011) and find a large degree of altruism but also substantial heterogeneity across physicians. With my experiment, I extend the analysis of heterogeneity in physician behavior and the results suggest that physician altruism varies not only between physicians but also systematically within them. This is in line with the concept of conditional altruism, i.e., that the degree of altruism toward the patient varies across patients with different medical needs.

Data from the experiment show that many physicians behave altruistically toward their patients but also that the degree of altruism varies across patient type, which indicates that physicians condition their altruism on certain characteristics of the patients. This variation across patients is mostly described by concerns for severity of illness, i.e., patients with more severe illnesses are given better medical care than otherwise identical patients with less urgent medical conditions. There are more selfish physicians

---

[13] See also, e.g., Woodward and Warren-Boulton (1984), Chalkley and Malcomson (1998), and Ma and Riordan (2002). Galizzi et al. (2015) provides a recent overview.

but also more purely altruistic physicians under fee-for-service than under capitation. Interestingly, even though the overall level of medical treatment differs substantially between capitation and fee-for-service payment, there is no significant difference in the distribution of physician types based on conditional altruism. The results also show a strong effect of payment system on physicians' provision behavior in general, consistent with previous studies using a similar design. The present paper is the first to show that the results extend into the domains of risk and uncertainty in the effect of medical treatment on patient health. While the overall pattern of provision behavior is consistent across domains, detailed analysis using subjects' generic risk and uncertainty preferences reveals individual heterogeneity in their reactions to the introduction of risk and uncertainty in patient health. This heterogeneity seems to depend on the payment system: There is a significant correlation between degree of risk aversion and response to patient risk and uncertainty under capitation but not under fee-for-service.

At a broader level, this paper relates to the discussion about using more detailed and fine-grained incentives in health care, such as performance-based payments targeting best practices and health outcomes for specific types of illnesses. Along these lines, a natural development would be to consider risk-differentiated payment systems in health care. For example, if physicians making decisions about risky treatments become overly cautious under certain types of payment systems, these types of incentives should be avoided in such situations. In the experiment, fee-for-service seems to be the marginally better payment system in this respect, since there is less heterogeneity in the effect of patient risk on physician behavior compared with the capitation payment system. One interpretation of this is that the monetary incentives under fee-for-service are stronger (perhaps they are more salient) than under capitation, thus leaving less room for decision making modulated by individual risk preference and other factors. However, since the magnitude of the difference in the effect of risk and uncertainty between capitation and fee-for-service is quite small, the overall implication for the organization of health care is that there is probably not much to gain from involving aspects of patient risk in the design of physician payment systems. Along these lines, an interesting aspect to consider in future research is the effect of the payment system on physician behavior in situations where more serious health outcomes are possible but also relatively unlikely to occur.

Medical need is a key aspect in selecting socially preferred allocations of scarce medical resources, in normative analysis and also in health policy. In discussing how physician behavior and altruism toward the patients correspond to different principles for priority

setting based on medical need, the capacity-to-benefit principle can be seen as the benchmark since it is directly connected to the marginal health benefit of the patient.[14] In light of this, it is interesting to note that a majority of physicians in the experiment do not behave in accordance with this ethical principle, but rather seem to factor in other aspects of medical need as well, such as the severity of illness. The popularity of this principle is in line with the public view in many countries, including France, the Netherlands, Norway, Sweden, and the UK, that health improvements are more valuable for people with more severe illnesses (e.g., Dolan et al., 2005; Shah, 2009; Nord and Johansen, 2014). This is also reflected in health policy. For example, in the UK, the National Institute for Health and Clinical Excellence (NICE) Citizen Council reached the conclusion that the severity of illness should be taken into account when making decisions, alongside the already established criteria based on cost and clinical effectiveness (NICE, 2008).[15] Similarly, in the early 1990s the State of Oregon increased the coverage of the publicly funded Medicaid health program, and the expansion included an explicit focus on health care prioritization based strictly on cost effectiveness in order to contain costs. However, the public reactions were negative and the policy was later adjusted so that other factors, including the severity of illness, were allowed to influence prioritization as well (Ham, 1998; Tinghög, 2011).

For health policy, the main takeaway from the experiment is that payment systems seem to affect neither the distribution of medical care nor physician response to risk and uncertainty in the outcome of medical treatment for the patient. One aspect in favor of the capitation payment system is that it seems to result in fewer cases of physicians only interested in own income. This payment system also generated a greater proportion of physicians motivated by a mixture of own income and altruism, and these physicians were generally responsive to patients' medical need in their treatment decisions. The heterogeneity within each payment system regarding how physicians' behavior corresponds to general principles for priority setting calls for careful consideration when designing payment systems in the future. An important aspect in this respect could be to develop clearer clinical guidelines, something that has been emphasized, e.g., in the UK. Another suggestion is to test mixed or performance-based payment systems in health care (e.g., Brosig-Koch et al., 2013, 2015a), and this could be further investigated in

---

[14] For example, in Ellis and McGuire (1986), the marginal rate of substitution between own income and health benefit for the patient determines how much medical treatment the physician wants to provide.

[15] NICE provides advice and guidance for clinical and public health professionals in the UK. A citizen council was established by NICE in 2002 in order to "provide advice about the social values that should underpin the Institute's guidance" (NICE, 2008). The council has met and discussed several different issues and in 2008 they met to discuss the case of severity of illness.

future research.

Acknowledgements

References

Arrow, K.J., 1963. Uncertainty and the welfare economics of medical care. American Economic Review 53 (5), 941–973.

Bénabou, R., Tirole, J., 2003. Intrinsic and extrinsic motivation. Review of Economic Studies 70 (3), 489–520.

Berger, L., Bleichrodt, H., Eeckhoudt, L., 2013. Treatment decisions under ambiguity. Journal of Health Economics 32 (3), 559–569.

Brosig-Koch, J., Hennig-Schmidt, H., Kairies, N., Wiesen, D., 2013. How effective are pay-for-performance incentives for physicians? A laboratory experiment. Ruhr Economic Papers 413.

Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., Wiesen, D., 2015a. The effects of introducing mixed payment systems for physicians – experimental evidence. Health Economics, forthcoming.

Brosig-Koch, J., Kairies-Schwarz, N., Kokot, J., 2015b. Sorting into physician payment schemes – a laboratory experiment. Ruhr Economic Papers 529.

Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., Wiesen, D., 2015c. Using artefactual field and lab experiments to investigate how fee-for-service and capitation affect medical service provision. Journal of Economic Behavior and Organization, forthcoming.

Camerer, C., Weber, M., 1992. Recent developments in modeling preferences: uncertainty and ambiguity. Journal of Risk and Uncertainty 5 (4), 325–370.

Chalkley, M., Malcomson, J., 1998. Contracting for health services when patient demand does not reflect quality. Journal of Health Economics 17 (1), 1–19.

Cookson, R., Dolan, P., 2000. Principles of justice in health care rationing. Journal of Medical Ethics 26 (5), 323–329.

Culyer, A.J., Wagstaff, A., 1993. Equity and equality in health and health care. Journal of Health Economics 12 (4), 431–457

Cutler, D., 2002. Equality, efficiency, and market fundamentals: the dynamics of international medical-care reform. Journal of Economic Literature 40 (3), 881–906.

Deci, E.L., 1971. Effects of externally mediated rewards on intrinsic motivation. Journal

of Personality and Social Psychology 18 (1), 105–115.

Deci, E. L., Ryan, R. M., 1985. Intrinsic Motivation and Self-Determination in Human Behavior, Plenum, New York.

Devlin, R.A., Sarma, S., 2008. Do physician remuneration schemes matter? The case of Canadian family physicians. Journal of Health Economics 27 (5), 1168–1181.

Dolan, P., Olsen, J.A., 2001. Equity in health: the importance of different health streams. Journal of Health Economics 20 (5), 823–834.

Dolan, P., Shaw, R., Tsuchiya, A., Williams, A., 2005. QALY maximization and people's preferences: a methodological review of the literature. Health Economics 14 (2), 197–208.

Dranove, D., Wehner, P., 1994. Physician-induced demand for childbirths. Journal of Health Economics 13 (1), 61–73.

Ellsberg, D., 1961. Risk, ambiguity, and the Savage axioms. Quarterly Journal of Economics 75 (4), 643–669.

Ellis, R., McGuire, T., 1986. Provider behavior under prospective reimbursement: cost sharing and supply. Journal of Health Economics 5 (2), 129–151.

Falk, A., Heckman, J., 2009. Lab experiments are a major source of knowledge in the social sciences. Science 326, 535–538.

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

Frey, B.S., Oberholzer-Gee, F., 1997. The cost of price incentives: an empirical analysis of motivation crowding-out. American Economic Review, 87 (4): 746–755.

Fuchs, V., 1978. The supply of surgeons and the demand for operations. Journal of Human Resources 13, 35–56.

Galizzi, M., Tammi, T., Godager, G., Linnosmaa, I., Wiesen, D., 2015. Provider altruism in health economics. National Institute for Health and Welfare Discussion Paper 4/2015.

Gaynor, M., Gertler, P., 1995. Moral hazard and risk spreading in partnerships. RAND Journal of Economics 26 (4), 591–613.

Godager, G., Wiesen, D., 2013. Profit or patients' health benefit? Exploring the heterogeneity in physician altruism. Journal of Health Economics 32 (6), 1105–1116.

Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. Journal of the Economic Science Association 1 (1), 114–125.

Gruber, J., Owings, M., 1996. Physician financial incentives and cesarean section delivery. RAND Journal of Economics 27 (1), 99–123.

Ham, C., 1998. Retracing the Oregon trail: the experience of rationing and the Oregon health plan. British Medical Journal 316, 1965–1969.

Hennig-Schmidt, H., Selten, R., Wiesen, D., 2011. How payment systems affect physicians' provision behavior–an experimental investigation. Journal of Health Economics 30 (4), 637–646.

Hennig-Schmidt, H., Wiesen, D., 2014. Other-regarding behavior and motivation in health care provision: an experiment with medical and non-medical students. Social Science and Medicine 108, 156–165.

Keser, C., Peterle, E., Schnitzler, C., 2014. Money talks: paying physicians for performance. Center for European Governance and Economic Development Research Discussion Paper 173.

Krasnik, A., Groenewegen, P.P., Pedersen, P.A., Scholten, P., Mooney, G., Gottschau, A., Flierman, H.A., Damsgaard, M., 1990. Changing remuneration systems: effects on activity in general practice. British Medical Journal 300, 1698-1701.

Ma, C.A., Riordan, M.H., 2002. Health insurance, moral hazard and managed care. Journal of Health Economics and Management Strategy 11 (1), 81–107.

McGuire, T., 2000. Physician agency, in: Culyer, A., Newhouse, J. (Eds.), Handbook of Health Economics. North Holland, Amsterdam.

Newhouse, J., 1992. Medical care costs: how much welfare loss? Journal of Economic Perspectives 6 (3), 3–21.

NICE, 2008. Quality adjusted life years (QALYs) and the severity of illness. NICE Citizens Council Meeting Report 10, http://www.nice.org.uk/Media/Default/Get-involved/Citizens-Council/Reports/CCReport10QALYSeverity.pdf

Nord, E., Johansen, R., 2014. Concerns for severity in priority setting in health care: a

review of trade-off data in preference studies and implications for societal willingness to pay for a QALY. Health Policy 116 (2–3), 281–288.

Olsen, J.A., 1997. Theories of justice and their implications for priority setting in health care. Journal of Health Economics 16 (6), 625–639.

Sabik, L.M., Lie, R.K., 2008. Priority setting in health care: lessons from the experience of eight countries. International Journal for Equity in Health 7: 4.

Shah, K.K, 2009. Severity of illness and priority setting in healthcare: a review of the literature. Health Policy 93 (2–3), 77–84.

Shigeoka, H., Fushimi, K., 2014. Supplier-induced demand for newborn treatment: evidence from Japan. Journal of Health Economics 35, 162–178.

Sutter, M., Kocher, M.G., Glätzle-Rützler, D., Trautmann, S.T., 2013. Impatience and uncertainty: experimental decisions predict adolescents' field behavior. American Economic Review 103 (1), 510–531.

Tinghög, G., 2011. The Art of Saying No: The Economics and Ethics of Health Care Rationing. PhD thesis, Linköping University.

Williams, A., 1997. Intergenerational equity: an exploration of the 'fair innings' argument. Health Economics 6 (2), 117–132.

Williams, B., 1962. The idea of equality, in: Laslett, P., Runciman, W.G. (Eds.), Philosophy, Politics and Society. Basil Blackwell, Oxford.

Williams, A., Cookson, R., 2000. Equity in health, in: Culyer, A., Newhouse, J. (Eds.), Handbook of Health Economics. North Holland, Amsterdam.

Woodward, R., Warren-Boulton, F., 1984. Considering the effect of financial incentives and professional ethics on 'appropriate' medical care. Journal of Health Economics 3 (3), 223–237.

Appendix A: Disaggregation of results by health profile

Insert Tables A1–A5 about here

Appendix B: Instructions for the experiment

[Included below are the original instructions for the CAP treatment. Whenever there is a difference between the CAP and FFS treatments, this is noted in the text.]

Welcome to the experiment and thank you for participating!

You are participating in an economic experiment on decision behavior. Please read the instructions carefully before the experiment starts. These instructions are solely for your private information. It is prohibited to communicate with others taking part in this experiment during the course of the session. Should you have any questions, please ask us by raising your hand and you questions will be answered privately. If you violate this rule, we will have to exclude you from the experiment and from all payments.

You and the other participants will be asked to make decisions for which you can earn money in addition to the 4 euros you will receive for participating in the experiment. Your payoff will depend on the decisions you make. At the end of the experiment, your payoff will be converted to euros and paid to you in cash. During the experiment, all amounts are presented in the experimental currency Taler.

1 Taler = 1 euro.

The experiment will take about 90 minutes and consists of three parts. You will receive detailed instructions prior to each part. The parts of the experiment are completely unrelated, which means that the decisions made in one part will not affect your earnings in the other parts. The sum of earnings from the different parts will constitute your total earnings from the experiment (together with the 4 euro show-up fee).

Part One

Please read the following instructions carefully. We will approach you in about 10 minutes to answer any questions you may have. If you have questions at any time during the experiment, please <u>raise your hand</u> and we will come to you. Part one of the experiment consists of 15 rounds of decision situations (and at the end of part one we will ask you to answer a few general questions related to health care and medical decision making).

Decision situations

In each round, you will take on the role of a physician and decide on the quantity of medical treatment for a given illness in a patient. That is, you will determine the quantity of medical services you wish to provide to the patient for a given illness.

The illnesses differ in two dimensions: (i) the relationship between the quantity of medical treatment and the resulting health benefit and (ii) the certainty in health benefit resulting from medical treatment.

- For illnesses in category 1, there is no risk or uncertainty involved in the treatment. This means that you will always know the exact health benefit resulting from the provided medical treatment. Below is a screenshot of the

29

information provided about the quantity of the medical treatment and the patient benefit (for the time being please focus only on the white columns). For example, if you provide the patient with a quantity of 2 medical services, the resulting health benefit is 7, and if you provide the patient with a quantity of 4 medical services, the resulting health benefit is 9.

Example patient
Illness in category 1 (no risk or uncertainty)

Quantity of medical treatment	Your capitation [fee-for-service] payment (in Taler)	Your costs (in Taler)	Your profit (in Taler)	Patient benefit (in Taler)	Your selected quantity of medical treatment
0				5	Provide 0
1				6	Provide 1
2				7	Provide 2
3				8	Provide 3
4				9	Provide 4
5				10	Provide 5
6				9	Provide 6
7				8	Provide 7
8				7	Provide 8
9				6	Provide 9
10				5	Provide 10

You are a physician.
Which quantity of medical treatment do you want to provide? | Confirm |

- For illnesses in category 2, there is risk involved in the treatment. More precisely, there is a 50% probability the medical treatment you provide will result in a good outcome, i.e., a high health benefit, but there is also a 50% probability of a bad outcome, i.e., a low health benefit. Below is a screenshot of the information provided about the quantity of medical treatment and the patient benefit (for the time being please focus only on the white columns). For example, if you provide the patient with a quantity of 2 medical services, the resulting health benefit is 8 if good outcome and 6 if bad outcome. Thus, the outcome of the treatment is risky. To characterize this, think of a bag containing 10 balls where 5 balls are white and 5 balls are black (we will call it bag A). The good outcome is realized if you guess correctly the color of a ball that will be drawn from the bag. If you guess that a white ball will be drawn randomly from the bag and this is what happens, then there is a good outcome for the patient, and consequently, if you provide the patient with a quantity of 2 medical services, the resulting health benefit will be 8. If you guess a white ball but a black ball is drawn, then there is a bad outcome for the patient, and consequently, if you provide the patient with a quantity of 2 medical services, the resulting health benefit will be 6. Similarly, if you provide the patient with a quantity of 4 medical services, the resulting health benefit will be 10 if good outcome and 8 if bad outcome.

Example patient
Illness in category 2 (risk)

Quantity of medical treatment	Your capitation [fee-for-service] payment (in Taler)	Your costs (in Taler)	Your profit (in Taler)	Patient benefit (in Taler)		Your selected quantity of medical treatment
				Good	*Bad*	
0				6	4	Provide 0
1				7	5	Provide 1
2				8	6	Provide 2
3				9	7	Provide 3
4				10	8	Provide 4
5				11	9	Provide 5
6				10	8	Provide 6
7				9	7	Provide 7
8				8	6	Provide 8
9				7	5	Provide 9
10				6	4	Provide 10

You are a physician.
Which quantity of medical treatment do you want to provide?   Confirm

-Help-
Risk is involved in the medical treatment.
There will be a draw from a bag (A) of 10 balls where 5 balls are white and 5 balls are black.
- There is a good outcome for the patient if you guess correctly the color of the drawn ball.
- There is a bad outcome for the patient if you guess incorrectly the color of the drawn ball.

- For illnesses in category 3, there is uncertainty involved in the treatment. More specifically, there is an unknown probability that the medical treatment you provide will result in a good outcome, i.e., a high health benefit, but there is also an unknown probability that the treatment will result in a bad outcome, i.e., a low health benefit. Below is a screenshot of the information provided about the quantity of medical treatment and the patient benefit (for the time being please focus only on the white columns). For example, if you provide the patient with a quantity of 2 medical services, the resulting health benefit is 8 if good outcome and 6 if bad outcome. Thus, the outcome of the treatment is uncertain. To characterize this, think of a bag containing 10 balls. You know there are both white and black balls in the bag, but not how many of each (we will call it bag B). The good outcome is realized if you guess correctly the color of a ball drawn from the bag. If you guess that a white ball will be drawn randomly from the bag and this is what happens, then there is a good outcome for the patient, and consequently, if you provide the patient with a quantity of 2 medical services, the resulting health benefit will be 8. If you guess a white ball but a black ball is drawn, then there is a bad outcome for the patient, and consequently, if you provide the patient with a quantity of 2 medical services, the resulting health benefit will be 6. Similarly, if you provide the patient with a quantity of 4 medical services, the resulting health benefit will be 10 if good outcome and 8 if bad outcome.

31

Example patient
Illness in category 3 (uncertainty)

Quantity of medical treatment	Your capitation [fee-for-service] payment (in Taler)	Your costs (in Taler)	Your profit (in Taler)	Patient benefit (in Taler)		Your selected quantity of medical treatment
				Good	*Bad*	
0				6	4	Provide 0
1				7	5	Provide 1
2				8	6	Provide 2
3				9	7	Provide 3
4				10	8	Provide 4
5				11	9	Provide 5
6				10	8	Provide 6
7				9	7	Provide 7
8				8	6	Provide 8
9				7	5	Provide 9
10				6	4	Provide 10

You are a physician.
Which quantity of medical treatment do you want to provide?   Confirm

-Help-
Uncertainty is involved in the medical treatment.
There will be a draw from a bag (B) of 10 balls where there are white and black balls.
You don't know the composition of the number of white and black balls in the bag.
- There is a good outcome for the patient if you guess correctly the color of the drawn ball.
- There is a bad outcome for the patient if you guess incorrectly the color of the drawn ball.

Profit

In each round you receive a capitation [fee-for-service] remuneration for treating the patient. Your remuneration is irrespective of the quantity of medical treatment provided. [Your remuneration increases with the quantity of medical treatment provided.] You also incur costs for treating the patient and they depend on the quantity of services you provide. [You also incur costs for treating the patient, which also depend on the quantity of services provided.] Your profit for each decision is calculated by subtracting these costs from the capitation [fee-for-service] remuneration.

Every possible treatment quantity yields a benefit for the patient – contingent on illness. Hence, when deciding on medical services, you will determine not only your own profit but also the patient's benefit.

In each round, you will receive detailed information on your screen (see below) for the respective patient. The information will show your capitation [fee-for-service] remuneration, your costs and your profit, as well as the patient benefit, for each possible quantity of medical treatment. Below is a screenshot of the information provided. For example, let's say you are about to decide what quantity of medical treatment to provide for an illness with a risky outcome. If you provide the patient with a quantity of 2 [8] medical services, you profit will be 9.6 and the patient's health benefit 8 if good outcome and 6 if bad outcome. If you instead provide the patient with a quantity of 4 [6]

medical services, your profit will be 8.4 and the patient's health benefit 10 if good outcome and 8 if bad outcome.

Example patient
Illness in category 2 (risk)

Quantity of medical treatment	Your capitation [fee-for-service] payment (in Taler)	Your costs (in Taler)	Your profit (in Taler)	Patient benefit (in Taler)		Your selected quantity of medical treatment
				Good	*Bad*	
0	10 [0]	0.0	10.0 [0.0]	6	4	Provide 0
1	10 [2]	0.1	9.9 [1.9]	7	5	Provide 1
2	10 [4]	0.4	9.6 [3.6]	8	6	Provide 2
3	10 [6]	0.9	9.1 [5.1]	9	7	Provide 3
4	10 [8]	1.6	8.4 [6.4]	10	8	Provide 4
5	10 [10]	2.5	7.5 [7.5]	11	9	Provide 5
6	10 [12]	3.6	6.4 [8.4]	10	8	Provide 6
7	10 [14]	4.9	5.1 [9.1]	9	7	Provide 7
8	10 [16]	6.4	3.6 [9.6]	8	6	Provide 8
9	10 [18]	8.1	1.9 [9.9]	7	5	Provide 9
10	10 [20]	10.0	0.0 [10.0]	6	4	Provide 10

You are a physician.
Which quantity of medical treatment do you want to provide?   Confirm

-Help-
Risk is involved in the medical treatment.
There will be a draw from a bag (A) of 10 balls where 5 balls are white and 5 balls are black.
- There is a good outcome for the patient if you guess correctly the color of the drawn ball.
- There is a bad outcome for the patient if you guess incorrectly the color of the drawn ball.

At the end of the experiment, one of the 15 rounds in the first part of the experiment will be chosen randomly. Your profit in this round will be paid to you in cash. If the selected round involves a risky or uncertain outcome, a draw of balls will determine whether the outcome for the patient is good or bad, as explained. In case of the risky outcome, you will select a color, either black or white, and then one randomly selected participant will draw a ball from a bag containing 5 black and 5 white balls (bag A). In case of the uncertain outcome, you will select a color, either black or white, and then one randomly selected participant will draw a ball from a bag containing 10 balls with unknown proportions of black and white balls (bag B).

In this part of the experiment, no patients will be physically present in the laboratory. Yet, the patient benefit does accrue to a real patient, since the amount resulting from your decision will be transferred to Christoffel Blindenmission Deutschland e.V., an organization funding treatment of patients with eye cataracts.

The transfer of money to Christoffel Blindenmission Deutschland e.V. will be made after the experiment by the experimenter. An official note of confirmation will be published on the experimental laboratory's blackboard on July 6. You will be able to verify that the transferred amount is exactly the same as the total patient benefit (in euro) resulting from the decisions made by all participants in the randomly chosen

situation.

Comprehension questions

Prior to the decision rounds, we kindly ask you to answer a few comprehension questions. They are intended to help familiarize you with the decision situations. If you have any questions about this, please raise your hand. Part one of the experiment will begin once all participants have answered the comprehension questions correctly.

Part Two [handed out after the completion of part one]

Please read the following instructions carefully. We will approach you in about 5 minutes to answer any questions you may have. If you have questions at any time during the experiment, please raise your hand and we will come to you. Please remember that 1 Taler = 1 euro.

Decisions

Part 2 is composed of 2×20 individual decision situations. In the first set of decision situation, you will choose between either receiving a certain amount of money for sure or drawing a ball from bag C with a chance to win 5 Taler. An example is shown below. As you can see, you will be asked to make repeated choices between receiving a certain amount of money for sure and drawing a ball from bag C. You can also see that the guaranteed amount you can choose increases in each row. If you choose to draw a ball from bag C, the drawing will be conducted as follows. We will fill the bag with 5 black and 5 white balls. One randomly selected person will be chosen to draw a ball from the bag. Before this person draws the ball, you will pick a color, black or white. If your chosen color matches the randomly drawn ball, you will win 5 Taler (otherwise zero). You will be asked to make 20 choices. For example, for the first choice situation, you will decide whether to draw a ball from bag C and have the possibility of winning 5 Taler or receive 0.25 Taler for sure. If you prefer to draw from the bag, then check the box on the left; if you would rather receive 0.25 Taler for sure, then check the box on the right. As you can see, the amount in Taler increases in the right column. Once you have chosen the guaranteed amount, you should repeat this decision throughout the remaining choice situations. For example, if you choose to receive 0.25 Taler for sure in the first choice situation, instead of drawing a ball from bag C, then you will obviously also choose 0.5 Taler for sure over drawing a ball from bag C in the next choice situation.

	Part 2			
1	Draw from bag C	○	○	0.25 Tokens for sure
2	Draw from bag C	○	○	0.5 Tokens for sure
3	Draw from bag C	○	○	0.75 Tokens for sure
4	Draw from bag C	○	○	1 Token for sure
5	Draw from bag C	○	○	1.25 Tokens for sure
6	Draw from bag C	○	○	1.5 Tokens for sure
7	Draw from bag C	○	○	1.75 Tokens for sure
8	Draw from bag C	○	○	2 Tokens for sure
9	Draw from bag C	○	○	2.25 Tokens for sure
10	Draw from bag C	○	○	2.5 Tokens for sure
11	Draw from bag C	○	○	2.75 Tokens for sure
12	Draw from bag C	○	○	3 Tokens for sure
13	Draw from bag C	○	○	3.25 Tokens for sure
14	Draw from bag C	○	○	3.5 Tokens for sure
15	Draw from bag C	○	○	3.75 Tokens for sure
16	Draw from bag C	○	○	4 Tokens for sure
17	Draw from bag C	○	○	4.25 Tokens for sure
18	Draw from bag C	○	○	4.5 Tokens for sure
19	Draw from bag C	○	○	4.75 Tokens for sure
20	Draw from bag C	○	○	5 Tokens for sure
				Confirm

In the second set of decision situations, you will make similar decisions between receiving a certain amount of money for sure and drawing a ball from a bag – but now from bag D. Bag D will consist of 10 balls, but the number of black and white balls will be a secret; otherwise the choice situations are similar to those in the first set (the amount in Taler increases in the right column).

Payoffs

When you have made all 40 choices, one of the choices will be randomly selected to determine your payoff. If you chose to draw from the bag in the randomly selected choice, you will need to select a color. First you will be asked to indicate whether you select a black ball or a white ball in case the randomly selected choice involves a draw from bag C. Then you will be asked to do the same in case the randomly selected choice involves a draw from bag D. In the end of the experiment, one randomly selected person will draw one ball from bag C and one ball from bag D. The randomly selected choice that will determine your payoff will be shown on your screen.

Part Three [handed out after the completion of part two]

[This part is not used in the current paper]

Tables

Table 1. First part of the experiment (timeline of physician decision making).

	Health benefit is *Deterministic*	Health benefit is *Risky*	Health benefit is *Uncertain*
Patient:	1,2,3,4,5 (A–E)	6,7,8,9,10 (A–E)	11,12,13,14,15 (A–E)

Note: Health profile in parentheses (Patient 1 has health profile A, Patient 2 has profile B, and so on).

Table 2. Effect of payment system on medical treatment.

	Average deviation from optimal treatment	Average <u>absolute</u> deviation from optimal treatment	Proportion selfish decisions	Proportion optimal decisions
FFS ($N = 64$)	2.48 (1.79)	2.53 (1.73)	26.3%	24.7%
CAP ($N = 66$)	-2.46 (1.55)	2.56 (1.43)	16.3%	19.7%
P-value	$< 0.01^a$	0.82^a	0.12^b	0.34^b

Note: [a] Mann-Whitney U test. [b] T-test with errors clustered on individuals. Standard deviation in parentheses. Deviation from optimal treatment is the difference between chosen level of medical treatment and the optimal level of medical treatment for each patient. For each individual, this value is averaged over the 15 patients in order to create 64 independent observations in FFS and 66 independent observations in CAP for the Mann-Whitney U tests. The proportions in the last two columns are calculated based on the total number of decisions and thus $N = 960$ in FFS and $N = 990$ in CAP.

Table 3. Effect of risk and uncertainty on medical treatment.

		Deterministic	Risk	Uncertainty
Average <u>absolute</u> deviation from optimal treatment:	FFS ($N = 64$)	2.45 (1.76)	2.55 (1.73)	2.58 (1.77)
	CAP ($N = 66$)	2.52 (1.52)	2.58 (1.44)	2.58 (1.46)

Null hypothesis	*P-value*[*]
FFS: Deterministic = Risk	0.15
FFS: Deterministic = Uncertainty	0.04
CAP: Deterministic = Risk	0.91
CAP: Deterministic = Uncertainty	0.79

Note: *Wilcoxon signed-rank tests. Standard deviation in parentheses. Absolute deviation from optimal treatment is the absolute difference between chosen level of medical treatment and the optimal level of medical treatment for each patient. For each individual, this value is averaged over the five patients within each treatment (Deterministic, Risk, and Uncertainty) in order to create 64 independent observations in FFS and 66 independent observations in CAP for the Wilcoxon signed-rank tests.

Table 4. Share of subjects providing optimal treatment in FFS.

Health profile	*Treatment*		
	Deterministic	Risk	Uncertainty
A	43.8%	40.1%	39.1%
B	14.1%	9.4%[*]	10.9%
C	25.0%	22.0%	25.0%
D	18.8%	23.4%[*]	23.4%
E	25.0%	25.0%	25.0%

Note: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Wilcoxon signed-rank tests on the difference from Deterministic for each health profile and treatment (Risk and Uncertainty).

Table 5. Share of subjects providing optimal treatment in CAP.

Health profile	*Treatment*		
	Deterministic	Risk	Uncertainty
A	4.5%	6.1%	6.1%
B	50.0%	42.4%	39.4%[*]
C	12.1%	13.6%	13.6%
D	24.2%	13.6%[**]	16.7%[*]
E	21.2%	15.2%	16.7%

Note: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Wilcoxon signed-rank tests on the difference from Deterministic for each health profile and treatment (Risk and Uncertainty).

Table 6. Generic risk preference and medical treatment (linear regression).

Dependent variable: *Absolute deviation from optimal treatment*	FFS		CAP	
	(1)	(2)	(3)	(4)
Risk treatment	0.105 (0.058)[*]	0.258 (0.214)	0.013 (0.079)	-0.719 (0.307)[**]
Risk for self	-0.089 (0.087)	-0.081 (0.089)	-0.127 (0.055)[**]	-0.169 (0.057)[***]
Risk treatment × Risk for self	-	-0.02 (0.023)	-	0.083 (0.039)[**]
Control variables	YES	YES	YES	YES
Observations	630	630	640	640
- Clusters	63	63	64	64

Note: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Errors are clustered on individuals (within parentheses). *Risk treatment* is a dummy for patients for whom the outcome of medical treatment is risky (and the base category is patients for whom the outcome of medical treatment is deterministic). *Risk for self* is the first row where the subject chose the certain amount of money (over the risky lottery) in the individual choice task (10 = risk neutrality). Control variables: *age, gender,* whether the subject *smokes,* tries to keep a healthy *diet*, number of visits to the *doctor* in the last three and twelve months, *health status* and *political views* on a scale from left to right. A total of three subjects (in FFS) switched multiple times in at least one of the two individual choice tasks and are therefore removed from the analysis.

Table 7. Generic uncertainty preference and medical treatment (linear regression).

Dependent variable: *Absolute deviation from optimal treatment*	FFS		CAP	
	(1)	(2)	(3)	(4)
Uncertainty treatment	0.143 (0.075)*	-0.067 (0.260)	0.05 (0.091)	-0.345 (0.184)*
Uncertainty for self	-0.091 (0.067)	-0.096 (0.067)	-0.036 (0.073)	-0.060 (0.078)
Uncertainty treatment × Uncertainty for self	-	0.009 (0.032)	-	0.048 (0.022)**
Control variables	YES	YES	YES	YES
Observations	630	630	640	640
- Clusters	63	63	64	64

Note: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Errors are clustered on individuals (in parentheses). *Uncertainty treatment* is a dummy for patients for whom the outcome of medical treatment is uncertain (and the base category is patients for whom the outcome of medical treatment is deterministic). *Uncertainty for self* is the first row where the subject chose the certain amount of money (over the uncertain lottery) in the individual choice task (10 = uncertainty neutrality). Control variables: *age*, *gender*, whether the subject *smokes*, tries to keep a healthy *diet*, number of visits to the *doctor* in the last three and twelve months, *health status* and *political views* on a scale from left to right. A total of three subjects (in FFS) switched multiple times in at least one of the two individual choice tasks and are therefore removed from the analysis.

Table 8. Proportion of selfish and altruistic physicians when patient health is deterministic, by treatment and health profile.

	Selfish decision at least once	Selfish decision always	Optimal decision at least once	Optimal decision always
FFS ($N = 64$)				
Profiles: A,B,C,D,E	31.3%	20.3%	46.9%	10.9%
- Profiles: A,B	31.3%	21.9%	45.3%	12.5%
- Profiles: C,D,E	29.7%	20.3%	32.8%	17.2%
CAP ($N = 66$)				
Profiles: A,B,C,D,E	27.3%	10.6%	53.0%	3.0%
- Profiles: A,B	24.2%	15.2%	51.5%	3.0%
- Profiles: C,D,E	22.7%	10.6%	27.2%	10.6%

Table 9. Provision of medical care to patients from health profiles C and D.

Dependent variable: *Absolute deviation from optimal treatment.*	Profile C	Profile D	*P-value*[*]
FFS ($N = 64$)			
Deterministic	2.39 (1.93)	2.59 (1.85)	0.05
Risk	2.41 (1.87)	2.75 (1.98)	< 0.01
Uncertainty	2.41 (1.95)	2.83 (1.97)	< 0.01
CAP ($N = 66$)			
Deterministic	2.45 (1.63)	2.56 (1.80)	0.22
Risk	2.45 (1.59)	2.68 (1.67)	0.05
Uncertainty	2.55 (1.69)	2.53 (1.60)	0.82

Note: *Wilcoxon signed-rank test. Standard deviation in parentheses. The absolute deviation from optimal treatment is the absolute difference between chosen level of medical treatment and optimal level of medical treatment for the patient.

Table 10. Distribution of physician types.

	Full sample		Only mix	
	FFS ($N = 64$)	CAP ($N = 66$)	FFS ($N = 39$)	CAP ($N = 50$)
1. Selfish	20.3%	12.1%	-	-
2. Altruistic (optimal C and E)	18.8%	10.6%	-	-
3. Mix				
Principle				
- Equality ex-post	15.6%	10.6%	25.6%	13.7%
- Capacity to benefit	21.9%	30.3%	35.9%	39.2%
- Severity of illness	23.4%	36.4%	38.5%	47.1%
P-value[*]	0.32		0.23	

Note: *Kruskal-Wallis test of equality of proportions between FFS and CAP. The type distribution is based on physician decisions for patients with deterministic benefit functions from health profiles C and E: *selfish* = always income-maximizing choice (q=0 in CAP and q=10 in FFS), *altruistic* = optimal care to both C and E, *equality ex-post* = better care to C than E, *capacity to benefit* = worse care to C than E, and *severity of illness* = provide C and E with equal amounts of care.

Table A1. Effect of payment system on medical treatment (by health profiles).

	Average deviation from optimal treatment	Average absolute deviation from optimal treatment	Proportion selfish decisions	Proportion optimal decisions
Profile: A, B				
FFS ($N = 64$)	2.48 (1.76)	2.56 (1.69)	26.3%	26.3%
CAP ($N = 66$)	-2.55 (1.51)	2.64 (1.40)	17.4%	24.8%
P-value	< 0.01[a]	0.59[a]	0.18[b]	0.76[b]
Profile: C, D, E				
FFS ($N = 64$)	2.48 (1.83)	2.51 (1.79)	26.2%	23.6%
CAP ($N = 66$)	-2.40 (1.64)	2.50 (1.49)	15.5%	16.3%
P-value	< 0.01[a]	0.99[a]	0.09[b]	0.22[b]

Note: [a] Mann-Whitney U test. [b] T-test with errors clustered on individuals. Standard deviation in parentheses. Deviation from optimal treatment is the difference between chosen level of medical treatment and optimal level of medical treatment for each patient. For each individual, this value is averaged over the 15 patients in order to create 64 independent observations in FFS and 66 independent observations in CAP for the Mann-Whitney U tests. The proportions in the last two columns are calculated based on the total number of decisions and thus N = 384 in FFS and N = 396 in CAP in the top panel and N = 576 in FFS and N = 594 in CAP in the bottom panel.

Table A2. Effect of risk and uncertainty on medical treatment in FFS (by health profile).

Health profile	*Treatment*		
	Deterministic	Risk	Uncertainty
A	1.22 (1.29)	1.30 (1.26)	1.30 (1.24)
B	3.75 (2.41)	3.92 (2.40)	3.86 (2.45)
C	2.39 (1.93)	2.41 (1.87)	2.41 (1.95)
D	2.59 (1.85)	2.75 (1.98)	2.83 (1.97)[*]
E	2.30 (1.82)	2.38 (1.86)	2.50 (1.87)[*]

Note: *p < 0.1, **p < 0.05, ***p < 0.01. Wilcoxon signed-rank tests on the difference from Deterministic for each health profile and treatment (Risk and Uncertainty). Standard deviation in parentheses. Dependent variable is the absolute deviation from optimal treatment (the absolute difference between chosen level of medical treatment and the optimal level of medical treatment for each patient).

Table A3. Effect of risk and uncertainty on medical treatment in CAP (by health profile).

Health profile	*Treatment* Deterministic	Risk	Uncertainty
A	4.27 (2.10)	4.06 (2.04)	4.12 (1.93)*
B	1.05 (1.21)	1.17 (1.22)	1.21 (1.20)
C	2.45 (1.63)	2.45 (1.59)	2.55 (1.69)
D	2.56 (1.80)	2.68 (1.67)	2.53 (1.60)
E	2.24 (1.64)	2.55 (1.67)**	2.49 (1.60)*

Note: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Wilcoxon signed-rank tests on the difference from Deterministic for each health profile and treatment (Risk and Uncertainty). Standard deviation in parentheses. Dependent variable is the absolute deviation from optimal treatment (the absolute difference between chosen level of medical treatment and the optimal level of medical treatment for each patient).

Table A4. Generic risk preference and medical treatment (by health profile).

	AB (1)	AB (2)	CDE (3)	CDE (4)
FFS				
Risk	0.119 (0.098)	0.357 (0.316)	0.095 (0.058)	0.192 (0.268)
Risk for self	-0.086 (0.083)	-0.073 (0.087)	-0.091 (0.091)	-0.086 (0.093)
Risk × Risk for self		-0.025 (0.033)		-0.010 (0.028)
CAP				
Risk	-0.078 (0.104)	-1.187 (0.478)**	0.073 (0.093)	-0.407 (0.385)
Risk for self	-0.114 (0.056)**	-0.177 (0.063)***	-0.136 (0.058)**	-0.163 (0.058)***
Risk × Risk for self		0.126 (0.055)**		0.055 (0.046)

Note: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Linear regression, dependent variable is the absolute deviation from optimal treatment in deciding on the quantity of medical treatment for a patient. Errors are clustered on individuals (in parentheses). *Risk* is a dummy for patients for whom the outcome of medical treatment is risky (and the base category is patients for whom the outcome of medical treatment is deterministic). *Risk for self* is the first row where the subject chose the certain amount of money (over the risky lottery) in the individual choice task (10 = risk neutrality). Control variables: *age*, *gender*, whether the subject *smokes*, tries to keep a healthy *diet*, number of visits to the *doctor* within the last three and twelve months, *health status* and *political views* on a scale from left to right.

Table A5. Generic uncertainty preference and medical treatment (by health profile).

	AB (1)	AB (2)	CDE (3)	CDE (4)
FFS				
Unc.	0.095 (0.074)	0.122 (0.209)	0.175 (0.095)	0.030 (0.372)
Unc. for self	-0.098 (0.067)	-0.010 (0.066)	-0.087 (0.071)	-0.095 (0.070)
Unc. × Unc. for self		-0.003 (0.022)		0.017 (0.047)
CAP				
Unc.	-0.016 (0.110)	-0.195 (0.389)	0.093 (0.109)	-0.445 (0.144)***
Unc. for self	-0.039 (0.073)	-0.050 (0.089)	-0.034 (0.075)	-0.067 (0.074)
Unc. × Unc. for self		0.022 (0.047)		0.066 (0.018)***

Note: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Linear regression; dependent variable is the absolute deviation from optimal treatment in deciding on the quantity of medical treatment for a patient. Errors are clustered on individuals (in parentheses). *Unc.* is a dummy for patients for whom the outcome of medical treatment is uncertain (and the base category is patients for whom the outcome of medical treatment is deterministic). *Unc. for self* is the first row where the subject chose the certain amount of money (over the uncertain lottery) in the individual choice task (10 = uncertain neutrality). Control variables: *age*, *gender*, whether the subject *smokes*, tries to keep a healthy *diet*, number of visits to the *doctor* within the last three and twelve months, *health status* and *political views* on a scale from left to right.

Figures

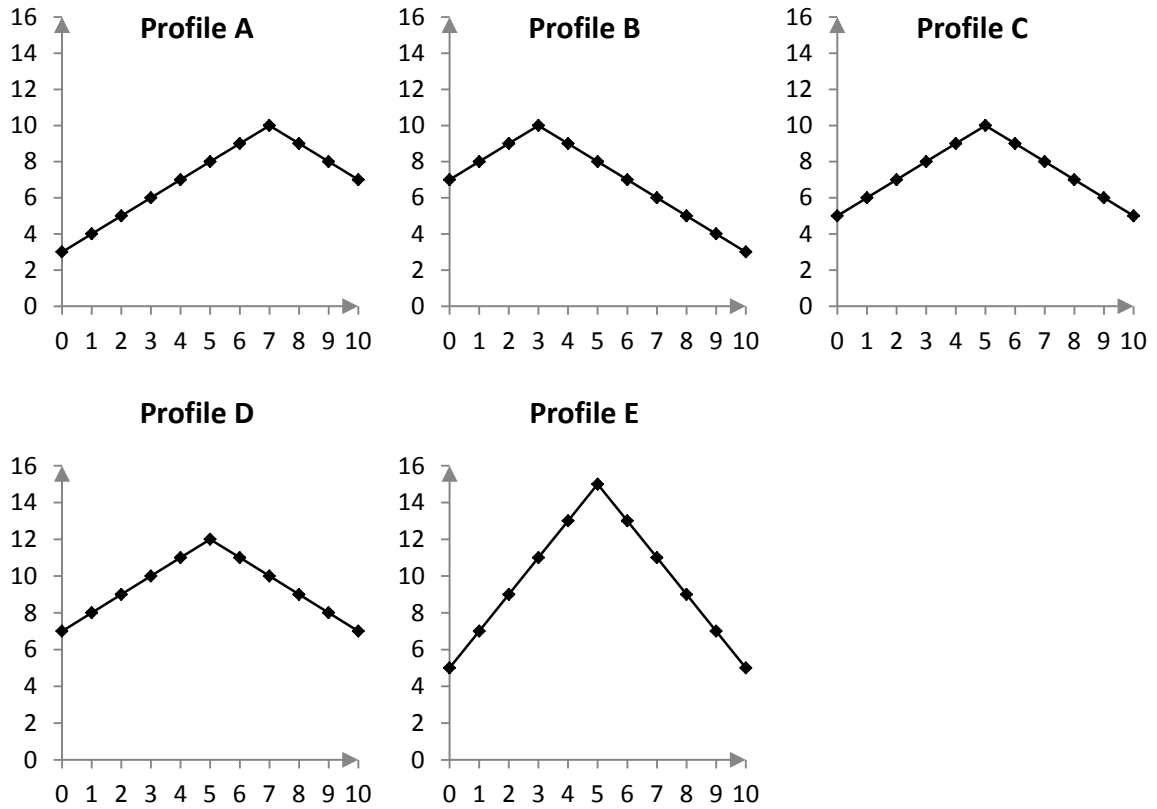**Figure 1.** Patient benefit functions (medical treatment (*q*) on the horizontal axes and health benefit on the vertical axes).

Figure 2. Average quantity of medical treatment per patient.