The Detrimental Effect of Group Size on Institution Formation

Julian Detemple[†] Michael Kosfeld[*]

April 22, 2025

Abstract

A solution for the provision of public goods is the formation of sanctioning institutions. While prior laboratory experiments document a causal effect of such institutions on cooperation in small groups, groups in the field are typically much larger. We test the causal effect of group size on institution formation and show that although individual preferences are similar, institution formation is about 45% more likely to fail as group size increases. Prior results thus do not generalize to larger groups. We document the voting rule and nonparticipants' increasing reluctance to change their mind as impeding factors in larger groups.

Keywords: Institution formation, group size, social dilemma
JEL Classification: C92, D02, H41

[†]Johannes-Gutenberg-University Mainz, Jakob-Welder-Weg 4, 55128 Mainz, Germany. E-mail: julian.detemple@uni-mainz.de. [*]Faculty of Economics and Business, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 4, 60323 Frankfurt am Main, Germany. E-mail: kosfeld@econ.uni-frankfurt.de. We thank seminar participants at Goethe University and Johannes-Gutenberg-Mainz University for helpful comments. We gratefully acknowledge support from the Experiment Center at the Leibniz Institute for Financial Research SAFE. IRB approval was obtained from the joint ethics committee of Goethe University and the University of Mainz prior to conducting the experiment. Declarations of interest: none.

1 Introduction

A key solution to collective action problems is the formation of institutions that change the rules of the game, e.g., through sanctions (Baland and Platteau, 1996; Ostrom, 1999). Examples include governance systems for communal resources, such as grazing lands in Switzerland (Ostrom, 1990) and forests in India (Agrawal and Yadama, 1997). However, often not all individuals are willing to participate in such institutions. Their formation thus depends on either rule enforcement by a higher authority or unanimous agreement among members to change the rules only for themselves, despite nonmembers also benefiting from collective action (Ostrom, 2005). This paper focuses on the latter case. Previous laboratory research has demonstrated that the opportunity to form such institutions enhances cooperation through the formation of both grand, i.e., institutions that involve everyone, and non-grand institutions (Dannenberg and Gallier, 2020). Yet, these findings are based on small groups, typically around four individuals, raising the question of how group size affects institution formation.

This paper addresses this question by examining the causal effect of increasing group size from four to eight and twelve individuals in a laboratory experiment. Importantly, the controlled environment allows to hold constant the profitability of an institution, i.e., the material incentives to form an institution, across group sizes. While group size does not affect preferences over which institution to form, we find that almost no institution is formed in groups with as few as eight players. We present evidence that this is due to the voting rule preventing the formation of non-grand institutions, coupled with nonparticipants' stronger reluctance to participate in a grand institution.

Our paper relates to two strands in the literature. First, our findings contribute to institution formation in lab experiments (e.g., Page, Putterman and Unel, 2005; Markussen, Putterman and Tyran, 2014; Kamei, Putterman and Tyran, 2015; Schmidt and Ockenfels, 2021). While groups with more than four individuals have been studied (e.g., Gürerk, Irlenbusch and Rockenbach, 2006; Kamei, Putterman and Tyran, 2019), we are the first to analyze the effect of group size through exogenous variation, documenting a negative effect.

Second, we contribute to how group size influences collective action in the field (e.g., Olson, 1965; Agrawal and Goyal, 2001; Poteete and Ostrom, 2004). In this literature, group size has a non-linear effect, with medium-sized groups exhibiting the highest success rate. While we study collective action in comparatively smaller groups in a laboratory

experiment, the controlled environment allows us to isolate the effect of group size, without affecting the material incentives to form an institution.

2 Institution Formation Game & Experimental Design

We adapt the n-person institution formation game analyzed in Okada (1993) and Kosfeld, Okada and Riedl (2009). The number of players n is common knowledge. As part of an institution, players commit themselves to contributing their endowment w to a linear public good. Players who do not commit themselves benefit from players' commitments and vice versa. To focus on the process of institution formation, commitments are enforced by assumption.[1]

Institutions are formed in three stages. First, players simultaneously and independently indicate their willingness to participate in an institution. We refer to players that (do not) indicate their willingness as *participants* (*nonparticipants*). Let s be the number of participants. If $s \geq 1$, the outcome of this stage is a proposed institution of size s. Second, all players are informed about the number of participants. Participants simultaneously and independently vote on the implementation of the proposed institution (if $s \geq 1$). Only if all participants unanimously vote in favor of implementation, the institution is formed and commitments become binding. Participants become *members*, nonparticipants *nonmembers*. Third, players are informed about the implementation outcome. If implementation was successful, nonmembers decide on their contributions, while members contribute their endowment. The implementation of the institution is costly and comes at a fixed cost c to each member.[2] If implementation was not successful, all players decide on their contribution. The marginal per capita return (MPCR) from contributing to the public good is δ/n with $0 < \delta/n < 1 < \delta$. By letting the MPCR depend on n, we deviate from Okada (1993) and Kosfeld, Okada and Riedl (2009), but this allows to hold the social, i.e., total, welfare effect of contributions to the public good constant relative to group size n. As we show in Appendix A.1, this ensures that the material incentives to implement an institution are constant across relative group sizes s/n.

Our experiment includes three treatments, with group sizes of $n = 4$ (*Treat-S*) as

[1] This setting has been shown to be equivalent to a sanctioning institution which enforces commitments via centralized sanctions for members (Kosfeld, Okada and Riedl, 2009).

[2] From an applied perspective, this cost may capture, e.g., resources needed to enforce commitments. For our analysis here, the cost is actually irrelevant as long as it not too large (which we ensure both in the model and in the experiment).

in Kosfeld, Okada and Riedl (2009), $n = 8$ (*Treat-M*), and $n = 12$ (*Treat-L*). Subjects participate in fixed groups for 20 rounds in a single treatment, with one randomly chosen round determining payoffs. Each subject receives an endowment of 10 points per round and institution implementation costs amount to 1 point. The parametrization of δ with $\delta = 2.4$ was chosen such that *Treat-S* is similar to Kosfeld, Okada and Riedl (2009).[3]

A total of 56, 120, and 132 participants participated in *Treat-S* (14 groups), *Treat-M* (15 groups), and *Treat-L* (11 groups), respectively. All data were collected from a student sample in 2020. Additional details on the experiment are provided in Appendix A.2.

3 Results

First, consider how group size affects overall institution formation.[4] Panel A in Table 1 displays the implementation rate of institutions across treatments, while Panel B shows average contributions to the public good per player, overall and depending on the implementation outcome. Group size affects both the formation of institutions and cooperation.

Table 1: Institution Formation & Cooperation

	Treat-S	*Treat-M*	*Treat-L*
A. Institution Formation			
Implementation Rate	0.41	0.15	0.11
B. Cooperation			
Contributions per Player	5.99	3.17	2.40
If implementation successful	9.16	9.24	8.65
If implementation failed	3.70	2.15	1.60

Notes: Group size is 4 (*Treat-S*), 8 (*Treat-M*), or 12 (*Treat-L*) players. All data are pooled on the treatment level.

The implementation rate drops from 0.41 in *Treat-S* to 0.15 in *Treat-M* ($p = 0.005$) and

[3] As an important side-result, *Treat-S* largely replicates the results from Kosfeld, Okada and Riedl (2009).

[4] Supplemental Appendix B.2 shows that all results are robust to accounting for learning. Data and replication code for all analyses are uploaded to an OSF repository and can be accessed via a private and anonymous link as part of the peer-review process (`https://osf.io/ck2eq/?view_only=8b203d966c1b4f2099e83b2883d18a9b`). We will enable public access to the OSF replication repository upon publication.

0.11 ($p = 0.004$), with no significant difference between *Treat-M* and *Treat-L*.[5] Cooperation also declines, from 5.99 points (out of 10 points) in *Treat-S* to 3.17 points in *Treat-M* ($p = 0.008$) and even 2.40 points in *Treat-L* ($p = 0.001$). Since contributions in case of a successful implementation do not change much, the significantly lower overall cooperation in larger groups can be traced to the lower implementation rate.

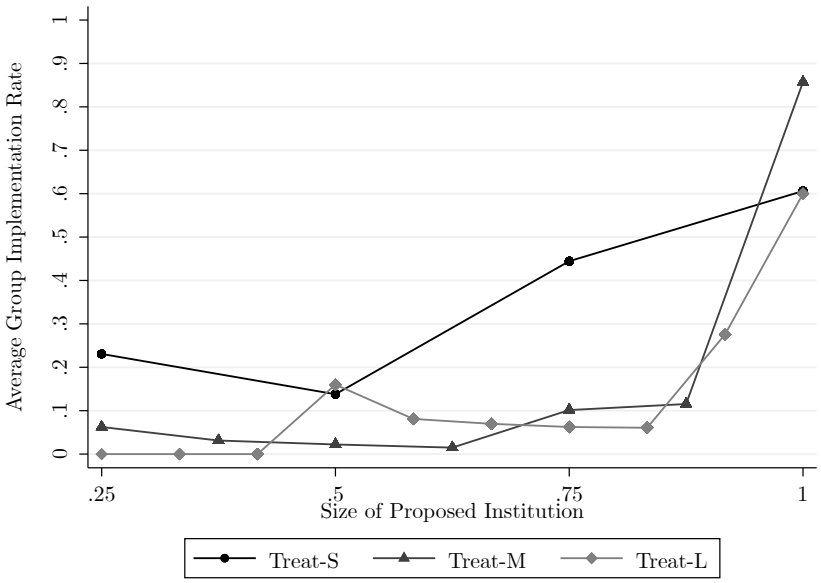

Figure 1: Implementation Rates of Proposed Institutions

Notes: Average implementation rates of institutions depending on the relative size of the proposed institution across all rounds. All data are pooled on the treatment level.

Second, potential explanations for the lower implementation rate are group-size effects on the implementation rate of different institution sizes (Figure 1) and the frequency with which such institutions are proposed (Table 2). Figure 1 shows that group size only impacts the implementation of so-called "almost-grand" institutions, i.e., non-grand institutions with at least 75% of the group participating. For example, when $n - 1$ players participate, implementation rates drop from 0.44 in *Treat-S* to 0.12 in *Treat-M* ($p = 0.034$) and 0.28 in *Treat-L* ($p = 0.057$). By contrast, the implementation of both grand institu-

[5]Tests are based on group averages. The respective p-values refer to pairwise two-sided Mann-Whitney-U tests with the exact p-values, unless noted otherwise.

tions and institutions with less than 75% of the group participating remains unaffected.[6] Moreover, Table 2 indicates that the lower implementation rate of almost-grand institutions is further exacerbated by fewer proposed grand institutions as group size increases. The share of grand institutions decreases from about 37% in *Treat-S* to 12% in *Treat-M* ($p = 0.011$) and even 2% in *Treat-L* ($p = 0.002$). Group size thus affects institution formation (and cooperation) by reducing both the frequency of proposed grand institutions and the implementation rate of almost-grand institutions.

Table 2: Proposed Institutions

	Treat-S	*Treat-M*	*Treat-L*
Share of Institution Type (in %)			
Grand Institution	36.94	11.78	2.27
Almost-Grand Institution	26.87	31.99	42.73
Other	36.19	56.23	55.00

Notes: A grand institution requires everyone to participate, i.e., $s/n = 1$. An almost-grand institution refers to $s/n \in [0.75, 1)$, with other referring to $s/n \in [0, 0.75)$. Group size (n) is 4 (*Treat-S*), 8 (*Treat-M*), or 12 (*Treat-L*) players. All data are pooled on the treatment level. See Supplemental Appendix B.2 for the unaggregated distribution.

The considerable share of nonparticipants suggests heterogeneity in participation preferences; the disagreement of participants about the implementation of almost-grand institutions suggests heterogeneity in preferences over institution size. We further explore these channels in the following. Consider the implementation of almost-grand institutions. Figure 2 plots the distribution of voting decisions on almost-grand institutions averaged on the player level, i.e., a measure for individual preferences over institution size, across treatments. Figure 2 highlights polarized preferences, with most participants willing to accept almost-grand institutions rather than no institution and around 20% strict naysayers (who always vote against an almost-grand institution). While group size does not affect the distribution of these preferences[7], it naturally increases the likelihood of having at least

[6] If at all, grand institutions are implemented more frequently in *Treat-M*. Differences are, however, not statistically significant ($p \geq 0.383$) and driven by outliers in *Treat-S* and *Treat-L*.

[7] The share of subjects who always vote against an almost-grand institutions and the share of subjects who always vote in favor of it are not statistically different across group sizes ($p \geq 0.331$, more details in Supplemental Appendix B.2).

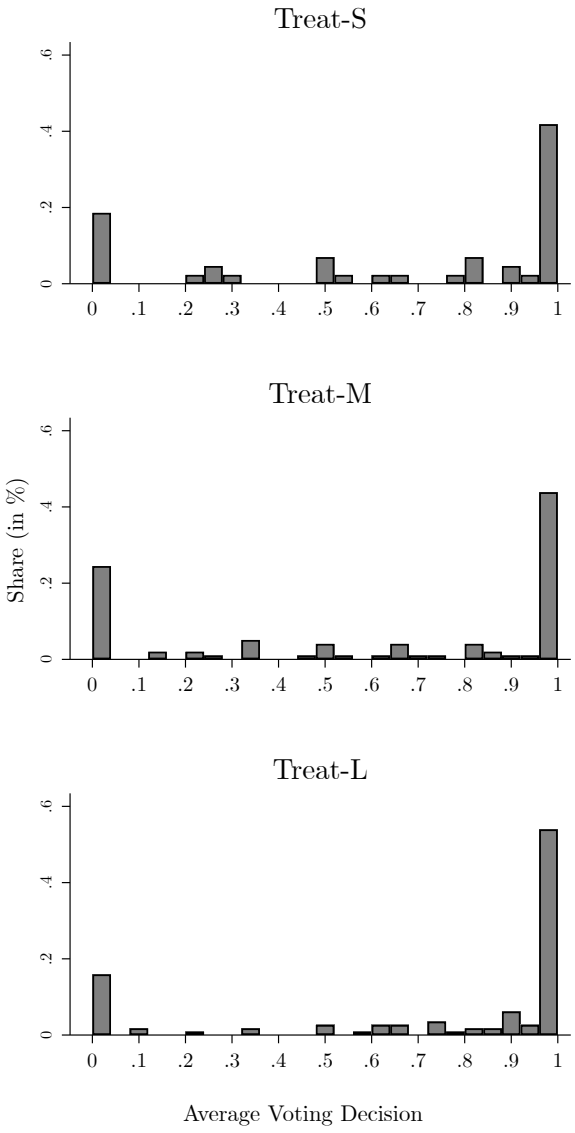

Figure 2: Distribution of Preferences over Almost-Grand Institutions

Notes: Distribution of preferences over the implementation of almost-grand institutions, i.e., the average voting decision on almost-grand institutions across all rounds. The average is computed as the average voting decision for each player who participated at least once in an almost-grand institution. All data are pooled on the treatment level.

one naysayer per group who blocks almost-grand institutions. Indeed, the share of groups with at least one such subject increases from 42% in *Treat-S* to 77% in *Treat-M* ($p = 0.082$, Fisher's one-sided exact test) and even 90% in *Treat-L* ($p = 0.026$, Fisher's one-sided exact test). The variation in these shares also explains the variation in implementation rates of almost-grand institutions both across and within treatments (Spearman's $\rho \leq -0.5609$ with $p \leq 0.046$). We can therefore attribute the lower implementation rate of almost-grand institutions to the unanimity rule and not a change in preferences.

Finally, consider that fewer grand institutions are proposed in larger groups. The data suggest that, across group sizes, participants very often continue to participate in the next round even when an almost-grand institution was rejected (86% in *Treat-S*, 89% in *Treat-M*, and 92% in *Treat-L*). The lower frequency of proposed grand institutions is thus due to nonparticipants not changing their mind. Miscoordination could explain this, with group size making it more difficult for multiple nonparticipants to collectively change their mind about commitments.[8] Miscoordination would imply that the likelihood of nonparticipants changing their mind should be comparable between large and small groups. This is not the case. When almost-grand institutions are rejected, the likelihood of a nonparticipant changing her mind in the following round significantly decreases from 45% in *Treat-S* to 28% in *Treat-M* and *Treat-L*. This finding is striking, considering that the payoff from a grand institution with full cooperation exceeds the payoff from no institution (cf. Appendix A.1). Consequently, increasing group size beyond four players not only severely hinders the formation of almost-grand institutions through the unanimity rule, but also prevents the rejection of almost-grand institutions from resolving the heterogeneity in participation preferences in favor of grand institutions.[9]

4 Conclusion

We study the effect of group size on the formation of institutions that require unanimity. We find that previous results do not generalize to group sizes beyond four players. Grand institutions are almost never proposed in larger groups, while almost-grand institutions are almost never implemented. The former is caused by nonparticipants who are more

[8] In contrast to *Treat-S*, *Treat-M* and *Treat-L* typically feature more than one nonparticipant in almost-grand institutions.

[9] At this point, we can only speculate about the underlying motivation but conjecture that group identification or social pressure might decline (Olson, 1965) and moral wiggle room might increase (Dana, Weber and Kuang, 2007) as groups become larger.

reluctant to change their mind about commitments, the latter stems from the unanimity rule coupled with similarly polarized preferences over institution size.

While simply requiring a majority decision seems like a solution at first glance, the lack of authority to enforce rules lies at the heart of many collective action problems. Future research may therefore study mechanisms for participants to reach a consensus on tolerating single or few nonparticipants and to motivate nonparticipants to change their mind (e.g., through activation of social norms).

References

Agrawal, Arun, and Gautam Yadama. 1997. "How Do Local Institutions Mediate Market and Population Pressures on Resources? Forest Panchayats in Kumaon, India." *Development and Change*, 28(3): 435–465.

Agrawal, Arun, and Sanjeev Goyal. 2001. "Group Size and Collective Action: Third-party Monitoring in Common-pool Resources." *Comparative Political Studies*, 34(1): 63–93.

Baland, Jean-Marie, and J. P. Platteau. 1996. *Halting Degradation of Natural Resources: Is There a Role for Rural Communities?* Oxford: Oxford University Press.

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments." *Journal of Behavioral and Experimental Finance*, 9: 88–97.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang. 2007. "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." *Economic Theory*, 33(1): 67–80.

Dannenberg, Astrid, and Carlo Gallier. 2020. "The Choice of Institutions to Solve Cooperation Problems: A Survey of Experimental Research." *Experimental Economics*, 23(3): 716–749.

Detemple, Julian. 2023. "Social Preferences: Why, How, and What to Measure." PhD diss. Goethe University, Frankfurt, Germany.

Fehr, E., and K. M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *The Quarterly Journal of Economics*, 114(3): 817–868.

Greiner, Ben. 2015. "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." *Journal of the Economic Science Association*, 1(1): 114–125.

Gürerk, Özgür, Bernd Irlenbusch, and Bettina Rockenbach. 2006. "The Competitive Advantage of Sanctioning Institutions." *Science*, 312(5770): 108–111.

Kamei, Kenju, Louis Putterman, and Jean-Robert Tyran. 2015. "State or Nature? Endogenous Formal versus Informal Sanctions in the Voluntary Provision of Public Goods." *Experimental Economics*, 18(1): 38–65.

Kamei, Kenju, Louis Putterman, and Jean-Robert Tyran. 2019. "Civic Engagement as a Second-Order Public Good: The Cooperative Underpinnings of the Accountable State." *CEPR Discussion Paper*, No. DP13985. Available at SSRN: https://ssrn.com/abstract=3464512.

Kosfeld, Michael, Akira Okada, and Arno Riedl. 2009. "Institution Formation in Public Goods Games." *The American Economic Review*, 99(4): 1335–1355.

Markussen, T., L. Putterman, and J.-R. Tyran. 2014. "Self-Organization for Collective Action: An Experimental Study of Voting on Sanction Regimes." *The Review of Economic Studies*, 81(1): 301–324.

Okada, Akira. 1993. "The Possibility of Cooperation in an N-Person Prisoners' Dilemma with Institutional Arrangements." *Public Choice*, 77(3): 629–656.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups. Harvard Economic Studies*. 21. printing ed., Cambridge, Mass.: Harvard Univ. Press.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action. The Political Economy of Institutions and Decisions*, Cambridge ; New York:Cambridge University Press.

Ostrom, Elinor. 1999. "Coping With Tragedies Of The Commons." *Annual Review of Political Science*, 2(1): 493–535.

Ostrom, Elinor. 2005. *Understanding Institutional Diversity. Princeton Paperbacks*, Princeton University Press.

Page, Talbot, Louis Putterman, and Bulent Unel. 2005. "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency." *The Economic Journal*, 115(506): 1032–1053.

Poteete, Amy R., and Elinor Ostrom. 2004. "Heterogeneity, Group Size and Collective Action: The Role of Institutions in Forest Management." *Development and Change*, 35(3): 435–461.

Prissé, Benjamin, and Diego Jorrat. 2022. "Lab vs Online Experiments: No Differences." *Journal of Behavioral and Experimental Economics*, 100: 101910.

Schmidt, Klaus M., and Axel Ockenfels. 2021. "Focusing Climate Negotiations on a Uniform Common Commitment Can Promote Cooperation." *Proceedings of the National Academy of Sciences*, 118(11): e2013070118.

Snowberg, Erik, and Leeat Yariv. 2021. "Testing the Waters: Behavior across Participant Pools." *American Economic Review*, 111(2): 687–719.

A Appendix

A.1 Institution Formation Game

Here we provide a more detailed description of how the game design allows to hold constant the material incentives to form an institution of the same relative size.

Let us introduce some further notation for this. A player's payoff u_i depends on whether an institution is successfully implemented and if yes, her role in it. Each player i has endowment w and we refer to the player's contribution with g_i. Remember that the marginal per capita return (MPCR) from contributing to the public good is δ/n with $0 < \delta/n < 1 < \delta$. While it is socially efficient to contribute w, individual incentives are such that each player has a strictly dominant action to contribute zero in case no institution is implemented. Let S denote the set of members of the institution with $s = |S|$, and let $c > 0$ be the institution formation cost for each member.

If $S \neq \emptyset$, i.e., an institution is implemented,

$$u_i = \begin{cases} \frac{\delta}{n}\left(sw + \sum_{j \notin S} g_j\right) - c & \text{if } i \in S, \\ w - g_i + \frac{\delta}{n}\left(sw + \sum_{j \notin S} g_j\right) & \text{if } i \notin S. \end{cases} \tag{1}$$

If $S = \emptyset$, i.e., no institution is implemented,

$$u_i = w - g_i + \frac{\delta}{n}\sum_{j=1}^{n} g_j. \tag{2}$$

Equations (1) and (2) highlight that nonmembers of an institution benefit from members' contributions and do not have to bear the institution formation cost c. In addition, they are free in their contribution decisions. This allows them to free-ride on members' commitments, leading to payoff inequality between members and nonmembers. Finally, we assume that $\delta w - c > w$, i.e., it is socially optimal to form the grand institution with $s = n$.[1]

We now outline that the material incentives to form an institution do not depend on group size under the assumption that nonmembers have money-maximizing preferences. We call an institution *profitable* if members' payoff in an institution is higher than without

[1] Of course, the first-best outcome would be if no institution was formed and all players contributed w. However, this outcome is no equilibrium.

any institution. If an institution of size s is implemented, members contribute w and nonmembers contribute nothing due to $\delta/n < 1$. If no institution is implemented, nobody contributes and everybody earns w. An institution of size s is thus profitable to implement if

$$\frac{\delta}{n} sw - c > w. \tag{3}$$

Since $\delta w - c > w$, sufficiently large institutions are profitable. Notice that for a fixed δ (as well as w and c), the relative size of an institution s/n alone determines whether an institution is profitable to implement. This allows us to vary group size and compare institutions of the same relative size while holding constant the profitability and thus material incentives of these institutions.

Notice that we have assumed that nonmembers have money-maximizing preferences and do thus not contribute to the public good. But nonmembers might, for example, exhibit social preferences such as compassion that motivate them to contribute to the public good despite not participating in the institution. If contributions by nonmembers and the share of contributing nonmembers is not affected by group size, the material incentives to form an institution are still independent of group size. Moreover, it is important to keep in mind that the revealed preferences over almost-grand institutions do indeed not depend on group size in our experiment (cf. Figure 2).

A full equilibrium analysis for both money-maximizing and social preferences – following Kosfeld, Okada and Riedl (2009), modeled as inequity-aversion based on Fehr and Schmidt (1999) – can be found in Detemple (2023) and is available upon request.

A.2 Experimental Details

Subjects played the institution formation game as described in Section 2. Before subjects could start with the experiment, they had to answer a set of control questions. See Supplemental Appendix B.1 for details. We piloted the experiment in August 2020 and conducted the sessions in August and September 2020 with oTree (Chen, Schonger and Wickens, 2016). Subjects were recruited using ORSEE (Greiner, 2015) from the subject pool of the Frankfurt Laboratory for Experimental Economic Research (FLEX) at Goethe University. Due to COVID-19 social distancing rules at that time, the experiment took place online.[2] All experimental sessions were embedded in a Zoom meeting with both cam-

[2]Snowberg and Yariv (2021) and Prissé and Jorrat (2022) find that subjects recruited from the same subject pool exhibit largely similar behavior in various experimental games regardless of whether the ex-

era and microphone turned off to protect subjects' privacy/identity. Subjects could only chat with the experimenter and not with each other. We sent out individualized tokens (6 random characters) by email 24 hours before a session started. Subjects were only allowed into the meeting with their names set to the individualized token, i.e., subjects did only see the random tokens of other subjects in the Zoom meeting. After all subjects had logged into the meeting, the experimenter gave a short verbal introduction and shared details on how to reach the experimenter in case of problems with the internet connection (see below). The experimenter then shared a link to the session on oTree and subjects had to enter their individualized token again. Subjects never got to know the tokens of others in the group they played the game in. Subjects were presented with the instructions and had to prove their understanding of the game by answering a similar set of control questions as in Kosfeld, Okada and Riedl (2009). If subjects answered a question incorrectly three times, they were informed about the right answer and could contact the experimenter in case of questions about why their answers were wrong (most did so). Moreover, subjects were also given time and opportunity to better understand the payoff consequences of different member/nonmember-constellations and average contribution by nonmembers, i.e., to learn which institutions are profitable to implement. For this, subjects could simulate the number of (non)members and the average contribution of nonmembers and were presented with the associated earnings of a member and nonmember on a separate "simulation"-page before the game began. Subjects were also provided with a summary of the relevant instructions underneath each decision. Subjects were paid via PayPal immediately after the session had ended and had agreed to this form of payment when signing up for the session.

Besides the possibility to message the experimenter on Zoom, subjects also received a mobile phone number which they could call, text, or message on Whatsapp in case of questions or problems of any sort (also after the session had ended in case of questions regarding the payment).

In total, 308 subjects completed the experiment.[3] On average, the experiment lasted 64 minutes and subjects earned 11.63€ with every point in the experiment worth 0.60€.

periment took place online or in the lab.

[3]Before the institution formation game, subjects also participated in another, unrelated task. However, there was no interaction among subjects in this task and subjects were informed about the outcome only after the overall experiment had ended.

B Supplemental Appendix

B.1 Instructions & Screenshots

The instructions are based on the instructions in Kosfeld, Okada and Riedl (2009). The following pages contain the English translation of the German instructions used in the experiment. For simplicity, we do not reproduce the original formatting which is very similar to the screenshots of the decisions screens.

Welcome to the second part of the experiment!

Welcome to the second part of this experiment. Please remember that you can also earn money in this part of the experiment. How much money you earn depends on your decision and the decision of the other participants in this experiment. At the end of the experiment you will receive the money you have earned via PayPal.

It is not allowed to communicate with other participants in this experiment in any way!

If you have a question, please contact us via the chat functionality in Zoom. We will then answer your question. When you click on the "Next"-button, we will explain the instructions for the second part of the experiment in detail. A summary of the relevant instructions will also be displayed together with every decision you will make in this experiment. In this experiment you will still make *all* decisions on your PC. You will still play with "points" in the second part of the experiment. For every point you will receive 0.60€ at the end of the experiment.

General instructions for the second part

Please read the following instructions for the second part very carefully! We will ask you some questions afterwards to check whether you understood the instructions.

The second part of the experiment consists of several rounds. You are still a member of the **same** group of participants as in the first part. Every group therefore has the **same 4 (8/12) participants** as in the first part of the experiment. But now you play with **all 3 (7/12) other group members together**. The group composition will not change. You

will still receive no information regarding the identity of the other group members – also not after the end of the experiment. The other participants in the experiment will also not receive any information on your identity. Any interaction in this experiment is therefore completely anonymous.

The second part of the experiment consists of **20 rounds**. All rounds are identical and consist of **4 phases**. All participants receive **10 points** at the beginning of each round. In the following, we will refer to these 10 points as **endowment**. Every round is structured as follows: in the **fourth** phase of each round, all group members receive information on the decisions which were taken during the previous phases of the round and the resulting earnings. In the **third** phase of each round, each of the 4 (8/12) group members decides how much of the endowment she wants to contribute to a common project and how much she wants to keep for herself. In the first **two phases** of each round, all group members decide whether they want to **commit themselves to contributing all of their endowment to the project**.

One of the 20 rounds will be **randomly selected** to be paid out at the end of experiment. When you will have completed the 20th round, you will be informed about which round was selected. Since you do not know which round will be selected by the random number generator, you should carefully consider your decisions in all rounds.

If a drop-out should happen in the second part of the experiment, you will receive the earnings from the first part of the experiment and a compensation for the second part. The amount of the compensation depends on how many rounds will have been played when the drop-out happens. If the drop-out happens before the end of the second round, it is not possible to randomly select a round. In this case you will receive 5 points as compensation. If the drop-out happens after the end of the second round, a previously played round will be randomly selected.

On the following pages all decisions in each phase of a round will be explained in detail. A summary of the instructions will also be made available every time you make a decision.

Instructions regarding your earnings

We will start by explaining how your earnings in a given round depend on your decisions in the third phase. On the following pages, we will explain the decisions in the first three phases. Please read the instructions very carefully.

In the third phase all 4 (8/12) group members need to decide how to allocate their endowment (10 points) between a personal account and a common project. Every point which you do not contribute to the common project is automatically transferred to your personal account. Your group members make exactly the same decision with the same rules governing their earnings.

Your earnings from the personal account: for every point which is transferred to your personal account, **you** receive 1 point and the **other** group members 0 points.

Your earnings from the common project: for every point which is contributed to the common project, **you** receive 0.6 (0.3/0.15) points and the **other** group members 0.6 (0.3/0.15) points. The sum of all earnings of all group members from the common project therefore increases by 2.4 points for every point you contribute to the common project. Your contribution to the project thus also increases the earnings of the other group members.

Your earnings from others' decisions: **regardless of how much you contribute to the project**, you receive 0.6 (0.3/0.15) points for every point contributed to the common project by another group member.

Potential costs: in addition to the earnings from your personal account and the common project, there can be costs which are deducted from your earnings. On the following page, we explain when these costs apply.

Your total earnings: your total earnings from this decision situation are the sum of the earnings from your personal account and the project, with any potential costs subtracted from it:

$$\text{Earnings} = (10 \text{ - your contribution to the project}) + 0.6 \times \text{the sum of project contributions of } \textbf{all} \text{ 4 group members - potential costs}$$

This decision situation is therefore similar to the first part of the experiment. The second part is different from the first part in that you are not matched with one randomly chosen person. Instead you are playing with all 3 (7/11) other group members simultaneously. Additionally, group members now have the opportunity to engage in commitments. In the following we explain how these commitments (and any potential costs) can matter.

Instructions regarding your decisions

Now we explain the decisions in all three phases.

First phase

In the first phase you decide whether you are **willing to commit yourself to contributing all 10 points** of your endowment to the **common project** in the third phase. Moreover, we ask you to estimate **how many of your 3 (7/11) other** group members are willing to commit themselves to contributing their endowment to the common project. When you are satisfied with your decisions, please click "Next". When every group member has made these decisions, the next phase starts.

Second phase

All group members who **were willing to commit to contributing all of their endowment** in the first phase now need to decide whether they **actually want to implement these commitments**. This happens as follows. If you were willing to commit yourself in phase 1, you now need to decide whether you **still want to commit** to contributing all of your endowment (10 points) to the common project. For this decision, you will receive information on the number of other group members who were willing to commit their endowment to the project in phase 3. Only if **all group members** who were **willing to commit themselves** in phase 1 are still willing to do so, the commitments **are implemented and become binding**. If, however, **at least one** group member – who was willing to commit herself – does not want to commit herself anymore, **all commitment** are **not implemented and do not become binding**.

Important: only group members who were willing to commit themselves in phase 1 make these implementation decisions. But all group members receive information on how many other group members were willing to commit themselves. **All** group members therefore receive **exactly the same** information, regardless of their decisions in the first phase.

If every group member who was willing to commit herself in phase 1 has made her decision, the next phase starts.

Third phase

In the third phase you have to choose your contribution to the project.

If the commitments became binding, the contribution decision will be automatically made for all group members who have committed themselves. These group members automatically contribute all of their endowment (10 points) to the project. Additionally, these group members also need to pay **extra costs**. These costs are 1 point and are paid

by every group member who has successfully committed herself.

If the commitment did not become binding, every group member can freely decide how many points of her endowment she wants to contribute to the project. In this case, there are **no** extra costs.

Important: the commitments (if they became binding) only apply to the group members who were willing to commit themselves in phase 1. Moreover, only those group members need to pay the extra costs. Group members who were not willing to commit themselves, can still freely choose their contribution to the project. When all group members have made this decision, the next phase starts.

Fourth phase

In the fourth (and last) phase of each round you receive information on how many points were contributed to the project in total. Moreover, we will show you your earnings of this round.

Please read these instructions very carefully. If you have questions, please write us in the Zoom chat. We will then respond to your question. When you have understood the instructions, you can click on "Next".

Simulation of earnings

[Please see the respective screenshot in Figure A.1.]

Control Questions

[All control questions contained a summary of the instructions and an online calculator.]

You will now see a few questions which we ask you to answer. These questions are meant to ensure that you have understood the instructions. You do not receive any money for answering these questions, but you have to answer them correctly to participate in the experiment. All questions are based on fictive examples. In these questions, we refer to the group members as A, B, C, and D (A, B, C, D, E, F, G, and H; A, B, C, D, E, F, G, H, I, J, K, and L).

Question 1: Suppose that nobody has committed herself to contributing all of her endowment to the project. Moreover, all 4 (8/12) group members contribute **0 points** to the project. What are the earnings of a single player in points?

Question 2: Suppose that nobody has committed herself to contributing all of her endowment to the project. Moreover, all 4 (8/12) group members contribute **10 points** to the project. What are the earnings of a single player in points?

Question 3a: Suppose that nobody has committed herself to contributing all of her endowment to the project. Player A (A and E; A, E, and I) contributes **2 points**, player B (B and F; B, F, and J) contributes **4 points**, player C (C and G; C, G, and K) contributes **5 points**, and player D (D and H; D, H, and L) contributes **9 points** to the common project. What are the earnings of player B in points?
Question 3b: What are the earnings of player D (H; L) in points?

Question 4a: Suppose that in the first phase half of all group members (players A, D; players A, D, E, H; players A, D, E, H, I, L) were willing to commit to contributing all of their endowment (10 points) to the project. Also suppose that these commitments have become **binding** in phase 2. In phase 3, the other group members (players B, D; players B, D, F, G; players B, D, F, G, J, K) contribute **0 points** points to the common project. How many points does player A (E/I) contribute to the project?
Question 4b: How high are the additional costs for player D in points? **Question 4c**: How high are the additional costs for player B (F/J) in points?

Question 5a: Suppose that in the first phase half of all group members (players A, D; players A, D, E, H; players A, D, E, H, I, L) were willing to commit to contributing all of their endowment (10 points) to the project. Also suppose that these commitments have become **binding** in phase 2. In phase 3, the other group members (players B, D; players B, D, F, G; players B, D, F, G, J, K) contribute **0 points** points to the common project. What are the earnings of player A in points?
Question 5b: What are the earnings of player C (G/K) in points?

Question 6a: Suppose that in the first phase half of all group members (players A, D; players A, D, E, H; players A, D, E, H, I, L) were willing to commit to contributing all of their endowment (10 points) to the project. Also suppose that these commitments have

not become binding in phase 2. In phase 3, all group members contribute **0 points** points to the common project. How high are the additional costs for player A (E/I) in points?

Question 6b: What are the earnings of player A in points?

Question 6c: What are the earnings of player B (F/J) in points?

[Please see Figures A.2 to A.7 for screenshots of the decision screens.]

Simulation of the decision situation

On this page you can simulate how the number of **successfully committed** group members and the (average) contribution of **not-committed** group members influences the earnings of the two types of group members respectively.
For this, please first choose the **number of successfully committed group members**:

○ 0 ○ 1 ◉ 2 ○ 3 ○ 4

Now please choose the **average contribution** of the **not-committed group members** to the common project in points:

○ 0 ◉ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

Important: in this experiment you can of course choose neither the number of successfully committed group members nor the average contribution of not-committed group members. This simulation exercise is just meant to make you acquainted with the overall decision situation.

The (average) earnings of a **successfully committed group member** would be **12,2** point(s).

The (average) earnings of a **not-committed group member** would be **22,2** point(s).

Please test how different examples influence the earnings of successfully committed and not-committed group members. When you have made yourself sufficiently acquainted with the decision situation, please click on "Next". In the following, we will ask you some questions which you need to answer correctly to start the second part of the experiment. A summary of the instructions which are relevant for each decision phase will always be provided on the decision screen.

Next

Summary of the instructions regarding your earnings

In each round your earnings are the sum of **two parts** and **potential costs** are deducted.

1. **Your earnings from the personal account** = number of points from your endowment which you keep for yourself = (10 points - your contribution to the project).
2. **Your earnings from the project** = 0.6 * sum of contributions of all 4 group members to the project.
3. **Potential costs**: if you **successfully commit** yourself to contributing all of your endowment to the project, you incur extra costs of 1 point. Otherwise there are no extra costs.

Summing up:

Your earnings = (10 points - your contribution to the project) + 0.6 * sum of contributions of **all 4 group members** to the project - potential costs

The earnings of each group member in your group are calculated in the same way.
Group members who have successfully committed themselves automatically contribute 10 points. Moreover, these group members also incur extra costs of 1 point. All other group members are free to decide how much they want to contribute to the project and do not incur any extra costs.

Figure A.1: Simulation Page

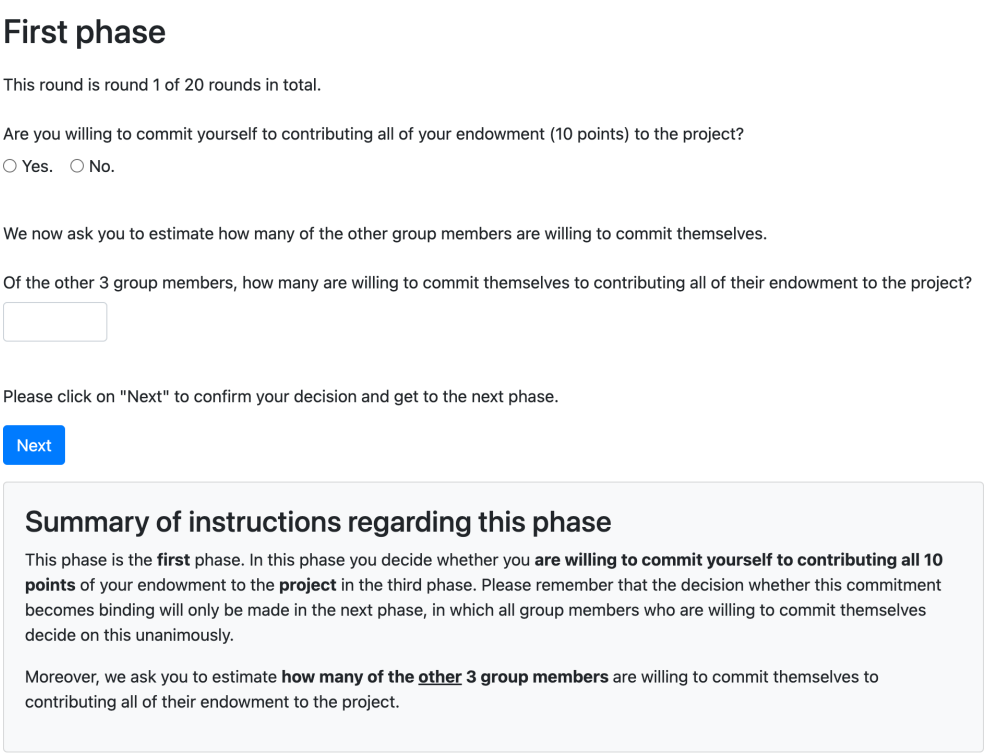

Figure A.2: First Phase (Participation Phase)

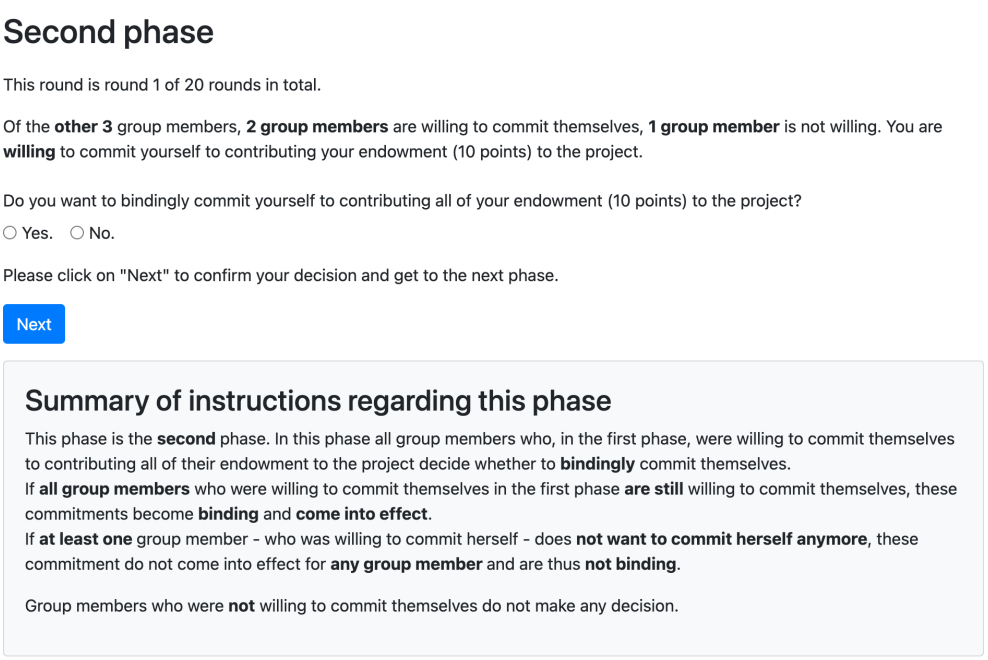

Figure A.3: Second Phase for Participants (Implementation Phase)

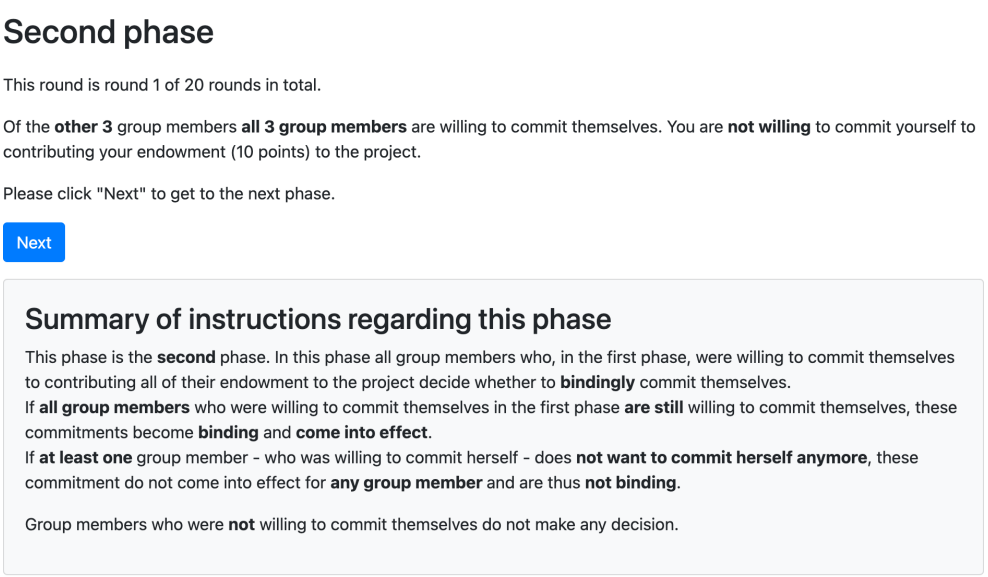

Figure A.4: Second Phase for Nonparticipants (Implementation Phase)

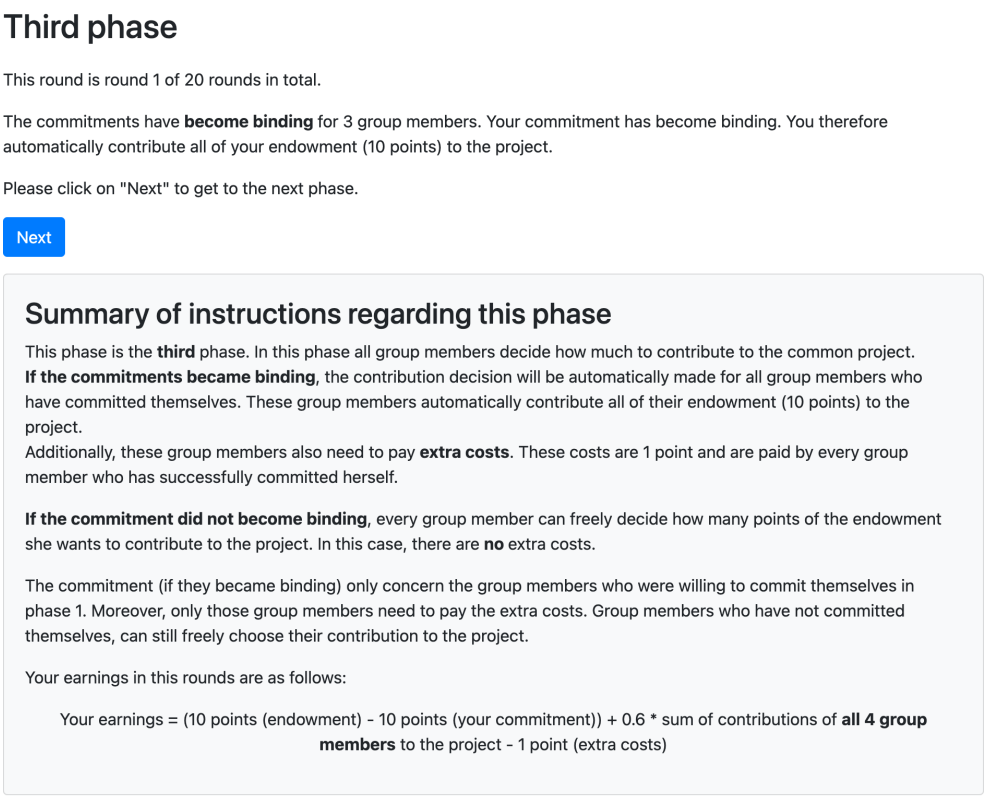

Figure A.5: Third Phase for Members (Contribution Phase)

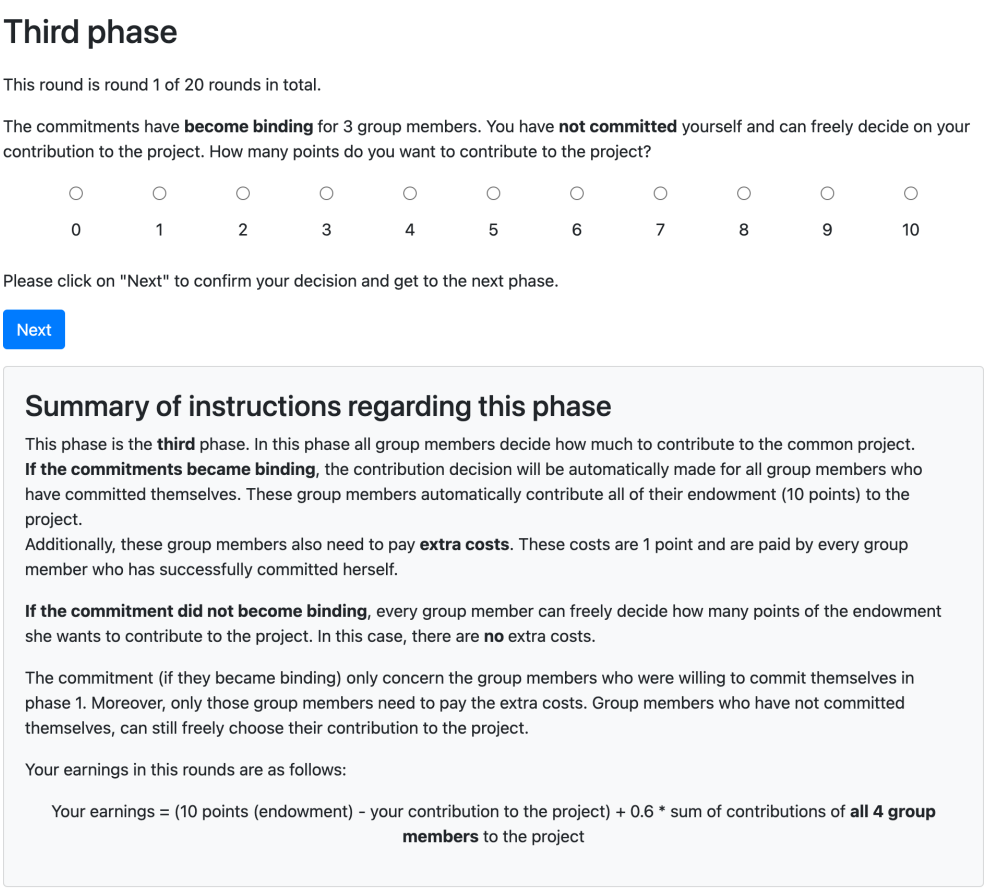

Figure A.6: Third Phase for Nonmembers (Contribution Phase)

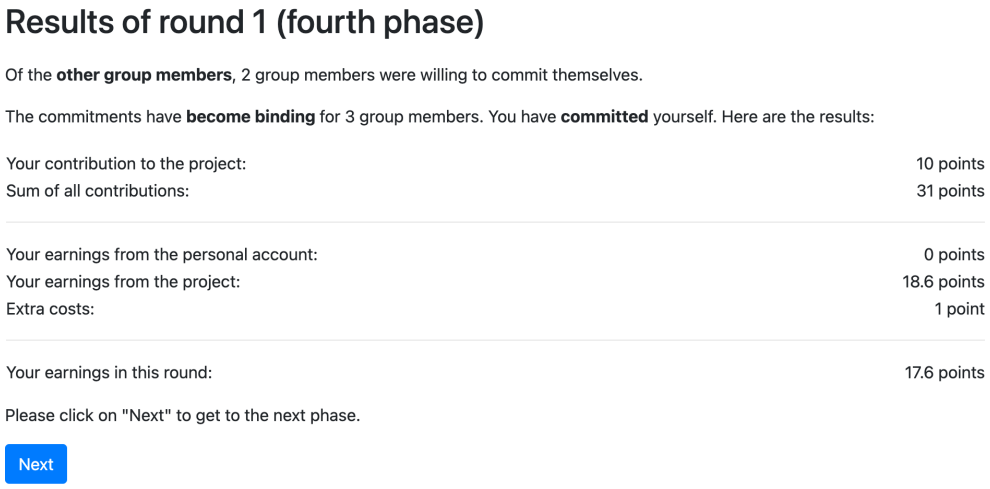

Figure A.7: Fourth Phase (Results)

B.2 Additional Analyses

B.2.1 Distribution of Proposed Institutions

Figure A.8 displays the full distribution of proposed institutions by institution size.

B.2.2 Group-size Effect on Preference Distribution

Figure 2 shows that the distribution of preferences over the implementation of almost-grand institution is polarized and very similar across all treatments. Statistical test indicate that there is no evidence in favor of a group-size effect on these distributions. First, the share of subjects who express a clear preference against almost-grand institutions (by always voting against the implementation) and the share of subjects who prefer an almost-grand institution over no institution (by always voting in favor of the implementation) are not statistically different across group sizes at conventional significance thresholds ($p \geq 0.331$). Second, this also holds when we consider subjects who sometimes vote in favor (or against) the implementation of an almost-grand institution: the shares of subjects who voted in favor of the implementation at most $x\% \in \{10, 20, 30\}$ or at least $y\% \in \{70, 80, 90, 100\}$ are not significantly different between *Treat-S* and any of the two larger group size treatments ($p \geq 0.295$). The same holds for the standard deviation in the individual average voting decision on almost-grand institutions ($p \geq 0.164$). Thus, group size does not seem to affect the distribution of individual preferences.

B.2.3 Robustness of Results to Learning

Since the institution formation game is a complex game and players might learn over time, results might change when we look at later rounds. However, all results are qualitatively similar when we only consider rounds 11-20 for our analyses.

First, Table A.1 replicates Table 1 for rounds 11-20. Overall implementation rates and average contributions per player are again significantly different when group size increases: implementation rates decrease from 0.44 in *Treat-S* to 0.18 in *Treat-M* ($p = 0.021$) and 0.15 in *Treat-L* ($p = 0.021$). Also in rounds 11-20, cooperation, measured as contributions per player, decreases when group size increases ($p = 0.004$ for the difference between *Treat-S* and *Treat-M* and 0.001 for the difference between *Treat-S* and *Treat-L*). Second, consider that these aggregate effects are driven by a lower implementation rate of almost-grand institutions, coupled with a lower frequency with which grand institutions are proposed.

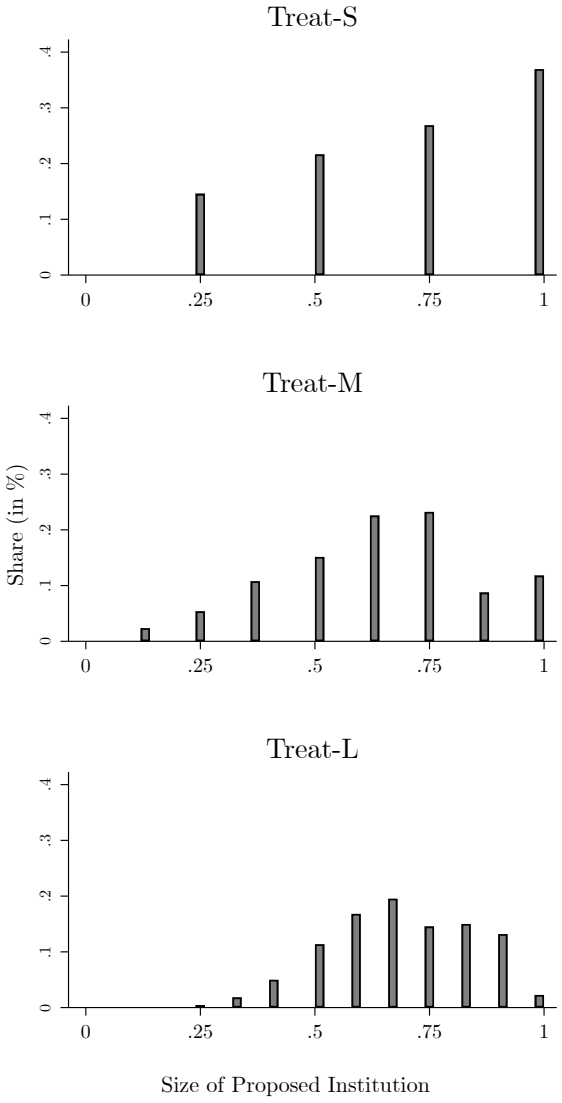

Figure A.8: Frequency of Proposed Institutions

Notes: Distribution of the relative size of proposed institutions over all rounds. All data are pooled on the treatment level.

Table A.1: Institution Formation & Cooperation (Rounds 11-20)

	Treat-S	*Treat-M*	*Treat-L*
A. Implementation Success			
Implementation Rate	0.44	0.18	0.15
B. Cooperation			
Contributions per Player	6.10	2.76	2.07
If implementation successful	9.33	9.39	8.74
If implementation failed	3.57	1.39	0.94

Notes: Group size is 4 (*Treat-S*), 8 (*Treat-M*), or 12 (*Treat-L*) players. All data are pooled on the treatment level for rounds 11-20 only.

Figure A.9 and Table A.2 replicate Figure 1 (implementation rates by the size of the proposed institution) and Table 2 (frequency of proposed institutions) for rounds 11-20.

Figure A.9 again shows a marked difference between small and larger groups in rounds 11-20. Institutions with three out of four subjects are more frequently implemented than institutions with seven (eleven) out of eight (twelve) subjects. By restricting the analysis to a subset of rounds, we do also lose some observations since not every group proposed an almost-grand institution in rounds 11-20. The statistical tests are thus based on fewer groups with the consequence that only the difference between *Treat-S* and *Treat-L* for the 75%-institution remains statistically significant ($p = 0.095$) at conventional levels ($p = 0.142$ for the difference in implementation rate of 75% institutions between *Treat-S* and *Treat-M*).[4] Table A.2 shows that also in rounds 11-20 grand institutions are proposed significantly less frequently in larger groups ($p = 0.012$ for the difference between *Treat-S* and *Treat-M* and $p = 0.006$ for the difference between *Treat-S* and *Treat-L*). Third, consider the individual channels. Figure A.10 replicates Figure 2 for rounds 11-20. It clearly shows that group size does not affect revealed preferences over almost-grand institutions. Similar to the main analysis, the great majority of subjects consistently vote in favor of the implementation of almost-grand institutions. Again, roughly 20% of subjects, however, always vote against the implementation of such institutions. Any difference in the share of

[4]To be specific, while the statistical tests for all all rounds are based on 12 groups for *Treat-S* as well as 13 and 9 groups for *Treat-M* and *Treat-L* respectively, the number of groups reduces to 9 (*Treat-S*) and 8 (*Treat-M* and *Treat-L*).

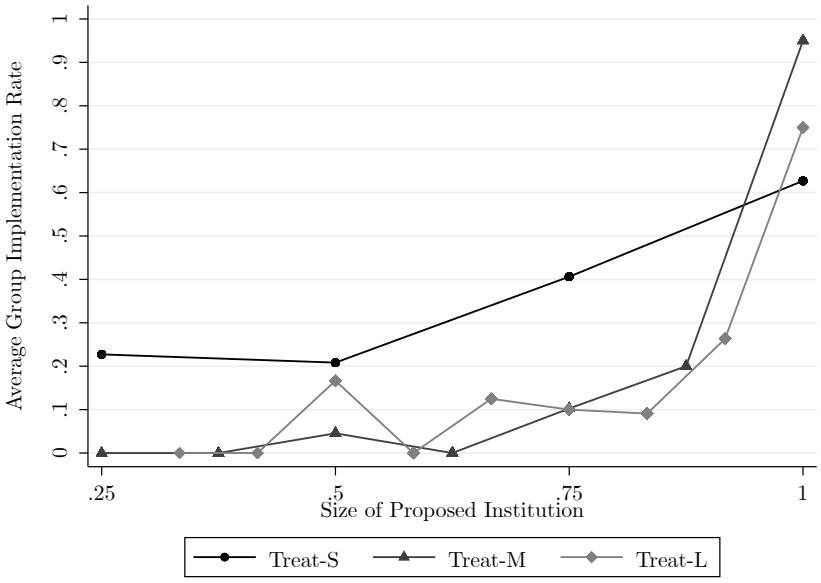

Figure A.9: Implementation Rates of Proposed Institutions (Rounds 11-20)

Notes: Average implementation rates of institutions depending on the relative size of the proposed institution in rounds 11-20. All data are pooled on the treatment level.

these subjects is not significantly different across group sizes ($p \geq 0.202$). Moreover, also the share of subjects who voted at most $x \in \{10, 20, 30\}$ or at least $y \in \{70, 80, 90, 100\}$ in favor of implementation of an almost-grand institution is not significantly different between *Treat-S* and any of the two larger group size treatments ($p \geq 0.202$). Consequently, group size does not seem to affect the distribution of preferences.

Moreover, we also document for rounds 11-20 that the share of groups with at least one subject who always votes against the implementation of almost-grand institutions strongly increases when increasing group size beyond four players. While only 44.44% of groups have one such player in *Treat-S*, this increases to 88.89% for both *Treat-M* and *Treat-L* ($p =$ 0.066). Similarly, having at least one player who always votes against the implementation of almost-grand institutions also explains the variation in the implementation rates of almost-grand institutions in rounds 11-20 (Spearman's $\rho =$ -0.7771 with $p = 0.000$).

Finally, consider the finding that the lower frequency of proposed grand institutions is driven by a lower willingness of nonparticipants to change their mind. Also in rounds 11-20, nonparticipants are less willing to participate in the following round when an almost-

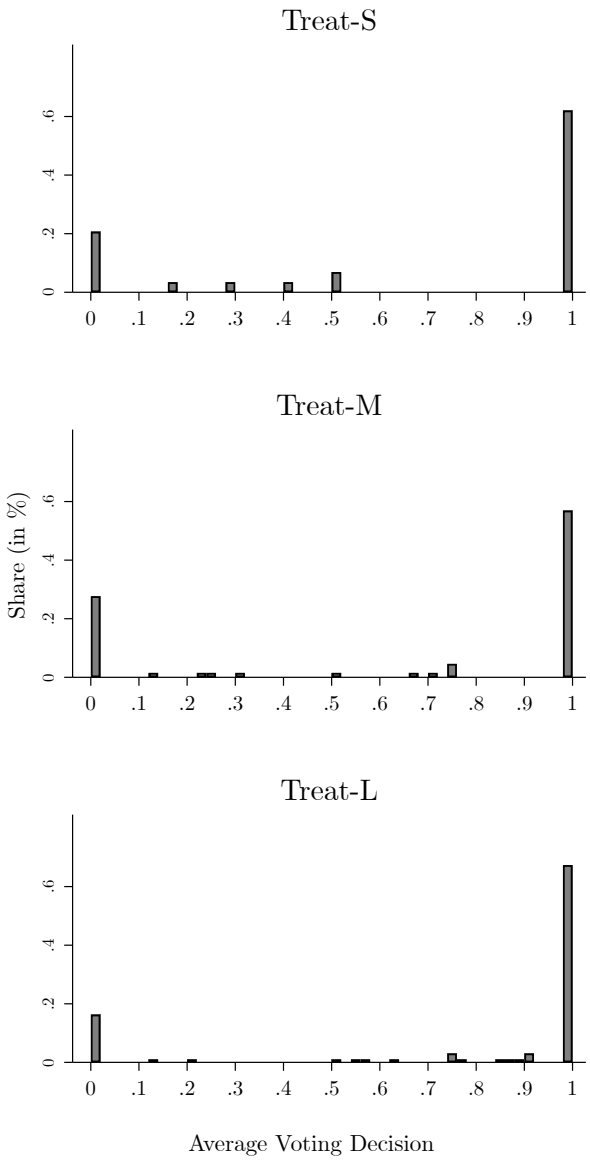

Figure A.10: Distribution of Preferences over Almost-Grand Institutions (Rounds 11-20)

Notes: Distribution of preferences over the implementation of almost-grand institutions, i.e., the average voting decision on almost-grand institutions in rounds 11-20. The average is computed as the average voting decision for each player who participated at least once in an almost-grand institution. All data are pooled on the treatment level.

Table A.2: Proposed Institutions (Rounds 11-20)

	Treat-S	*Treat-M*	*Treat-L*
Grand Institution	43.07	13.61	3.64
Almost-Grand Institution	23.36	33.33	55.45
Other	33.58	53.06	40.91

Notes: A grand institution requires everyone to participate, i.e., $s/n = 1$. An almost-grand institution refers to $s/n \in [0.75, 1)$, with other referring to $s/n \in [0, 0.75)$. Group size (n) is 4 (*Treat-S*), 8 (*Treat-M*), or 12 (*Treat-L*) players. All data are pooled on the treatment level for rounds 11-20 only.

grand institution was rejected (87% in *Treat-S*, 89% in *Treat-M*, and 94% in *Treat-L*). The likelihood that a nonparticipant participates in the following round decreases from 47.62% in *Treat-S* to 22.22% in *Treat-M* and 26.04% in *Treat-L*.