Discussion Papers
Department of Economics
University of Copenhagen

No. 19-07

Voting on Sanctioning Institutions in Open and Closed Communities: Experimental Evidence

Ramón Cobo-Reyes, Gabriel Katz, Thomas Markussen and Simone Meraglia

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark
Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (E)

Voting on Sanctioning Institutions in Open and Closed Communities: Experimental Evidence*

Ramón Cobo-Reyes[†], Gabriel Katz[‡], Thomas Markussen[§], Simone Meraglia[¶]

Abstract

We experimentally analyze the effect of endogenous group formation on the type of sanctioning institutions emerging in a society. We allocate subjects to one of two groups. Subjects play a repeated public goods game and vote on the sanctioning system (formal or informal) to be implemented in their group. We compare this environment to one in which subjects are allowed to (*i*) vote on the sanctioning system and (*ii*) move between groups. We find that the possibility of moving between groups leads to a larger proportion of subjects voting for formal sanctions. This result is mainly driven by subjects in groups with relatively high initial levels of contribution to the public good, who are more likely to vote for informal sanctions when groups are closed than when they are open.

Keywords: Sanctions, Cooperation, Group Formation, Voting, Experiment.

JEL Classification Numbers: C73, C91, C92, D72, H41.

*The authors are thankful to the University of Exeter BID Cluster and Independent Research Fund Denmark (project ID 4182-00163 FSE) for financially supporting this project. We thank participants at the 2018 Copenhagen Workshop on Endogenous Institutions in Social Dilemmas for comments. All remaining errors are ours.

[†]Department of Economics, American University of Sharjah. E-mail: rcobo@aus.edu

[‡]Department of Politics, University of Exeter. E-mail: G.Katz@exeter.ac.uk

[§]Department of Economics, University of Copenhagen. E-mail: thomas.markussen@econ.ku.dk

[¶]Department of Economics, University of Exeter. E-mail: s.meraglia@exeter.ac.uk.

1 Introduction

The presence and functioning of sanctioning institutions is key to overcoming free-riding problems (Ostrom, Walker and Gardner, 1992; Fehr and Gächter, 2000). Sanctioning institutions can take different forms, and most societies are involved in a dynamic process of institutional adaptation. A sizable literature has recently recognized the relevance of the endogenous choice of sanctioning institutions (e.g. Ertan, Page and Putterman, 2009; Sutter, Haigner and Kocher, 2010; Andreoni and Gee, 2012). In particular, increasing attention has been devoted to the choice between decentralized, peer-to-peer sanctions, and more centralized and formalized forms of punishment. Several studies find considerable support for peer-to-peer sanctions, even when centralized sanctions are theoretically expected to prevail (Markussen, Putterman and Tyran, 2014; Zhang et al., 2014; Kamei, Putterman and Tyran, 2015; Nicklisch, Grechenig and Thöni, 2016). This paper asks whether this result continues to hold when societies are open to entry and exit. In other words, we investigate whether open societies develop different sanctioning institutions than closed ones.

To the best of our knowledge, this is the first paper that studies how the formation of sanctioning institutions varies between open and closed societies. The topic is important in a number of contexts where mobility across group boundaries is changing. In developing countries, population growth, improvements in infrastructure and changing economic opportunities lead to increased internal migration, for example from rural to urban areas (Harris and Todaro, 1970; Deshingkar and Anderson , 2004). In large organizations such as companies, the increasing use of temporary employment (cf. the "gig economy") means that the composition of work teams changes at more rapid pace than before (De Stefano, 2016). And at the international level, wars, political oppression and gaps in economic opportunities lead many to migrate across borders. The experiment presented here contributes to our understanding of how increased mobility changes preferences between informal, horizontal institutions, and more centralized, formal arrangements, such as the rules and regulations

that a state or the management of a large organization may implement.[1]

There are two main approaches to studying the choice between sanctioning institutions in social dilemmas. In one strand of the literature, groups of fixed size and composition vote by ballot about which type of institution to use (e.g. Sutter, Haigner and Kocher, 2010; Andreoni and Gee, 2012; Markussen, Putterman and Tyran, 2014). Another cluster of experiments allows participants to "vote with their feet" by migrating to the institutional environment they prefer (e.g. Gürerk, Irlenbusch and Rockenbach, 2006, 2014; Nicklisch, Grechenig and Thöni, 2016; Zhang et al., 2014).[2] We combine these two approaches. In the baseline treatment, subjects are randomly allocated to two groups of 5 people to play a Public Goods Game (PGG). Subjects interact for 30 periods and groups' size and composition remain fixed for the entire duration of the experiment. Subjects vote every five periods on the sanctioning institutions – either formal or informal – to be implemented in their group. With informal sanctions, participants receive information about the contribution decisions of their peers and are given the opportunity to reduce the earnings of other group members, at a cost to themselves (cf. Fehr and Gächter, 2000). With formal sanctions, free riders are automatically punished. The formal sanctioning systems comes with a fixed cost, which must be paid by all group members. We compare the baseline with our main treatment, in which subjects not only vote about institutional choice, but can also (freely) move between groups (cf. Tiebout, 1956). In other words, we compare an environment where both "voice" (voting by ballot) and "exit" (voting with feet) strategies are available – i.e., "open societies" – to one where only the "voice" option exists – i.e., "closed societies" – (cf. Hirschman, 1970).

A priori, it is far from clear whether one should expect individuals to prefer more or less centralized forms of sanctioning institutions when groups' size and composition can vary.

[1] A relevant historical example is merchant guilds in Europe (Ogilvie, 2011; Dessi and Piccolo, 2016). We may view the guilds as a type of (relatively) informal, horizontal institution. Around 1800, as trade and mobility were expanding in the wake of the industrial revolution, many guilds were broken up. To a large extent, they were replaced with government regulation such as patent systems. Hence, this is an example of a move toward more formal, centralized institutions, which coincided with the emergence of a more dynamic, open society.

[2] As discussed below, exceptions include Robbett (2014) and Cobo-Reyes, Katz and Meraglia (2019).

From a theoretical point of view, Fehr and Schmidt (1999) show that if people have sufficiently strong social preferences (in particular, aversion to inequality), then multiple equilibria exist in public goods games with decentralized, informal sanctions (IS), some of them with high levels of cooperation and some with low levels. Reactions to opening borders may differ markedly, depending on whether groups have previously converged to cooperative or non-cooperative equilibria. On the one hand, if a group has established a cooperative equilibrium, allowing inward migration is a risk factor that may potentially lead the equilibrium to break down. Hence, groups that have established high levels of contributions under informal sanctions should vote for continued use of IS when their size and composition is fixed, but may prefer centralized, formal sanctions (FS) when migration is allowed, because the influx of new members potentially leads to the breakdown of cooperative equilibria. On the other hand, if cooperative norms have not been established, inward and outward migration could either provide an opportunity to move to a better equilibrium or it could make cooperation even harder to achieve. Therefore, we hypothesize that IS will lead to heterogeneous outcomes across groups, and that the outcomes of the initial phases of play with IS will have a strong impact on how the opportunity to move between groups affects voting patterns. To test this hypothesis, we fix groups during the initial phase of the experiment and impose IS (rather than FS) in all of them. After this initial phase, voting between IS and FS is introduced and the opportunity to move between groups is made available in one treatment.

We find that endogenous group formation dramatically affects institutional choice, but in different ways depending on groups' initial conditions. In particular, in our baseline treatment, subjects in groups with high average contributions to the public good in the initial phase of the experiment vote in favor of Informal Sanctions more than 80% of the time; this proportion decreases to less than 55% when subjects can move between groups. When subjects start out in groups with low average contributions to the public good in the initial phase of the experiment, the opposite results hold. In this case, 33% of the subjects vote for Informal Sanctions when groups are fixed, whereas more than 45% of the subjects

choose Informal Sanctions when they can move between groups. Results also show that, when given the opportunity, on average 20% of the subjects move between groups in each period. Subjects' group choice is mainly determined by the difference in contribution between the subject and other members of her group, and by the punishment she receives. We also find that although contributions are on average larger when Formal Sanctions are in place, this difference does not affect migration behavior and subjects do not tend to move systematically towards societies implementing Formal Sanctions.

Only a few other experimental studies on social dilemmas combine "voting by ballot" and "voting with feet" in the same experimental set-up. Robbett (2014) implements a non-linear public goods game and studies the choice of taxes, rather than sanctioning systems. She compares a treatment where only voting with feet is possible to one where both voting with feet and by ballots is allowed. In contrast, our treatments differ by the availability of the voting with feet option. This is how we distinguish "open" from "closed" societies. We focus on the choice of institutions, whereas Robbett is mainly interested in efficiency. Cobo-Reyes, Katz and Meraglia (2019) study the choice between formal sanctions and no sanctions at all in a public goods game but, like Robbett (2014), vary only the availability of the voting by ballot option, while voting with feet is always allowed. Hence, none of these papers compare voting about institutions in open and closed societies. Doing so is the main focus of this paper.

The paper proceeds as follows. Section 2 describes the experimental design. Section 3 reports and discusses our findings. Section 4 concludes.

2 Experimental Design and Procedures

We conducted laboratory experiments at the University of Exeter between October 2017 and February 2018. Subjects were mainly students of economics, business administration, and engineering, but other disciplines were also represented. We ran two different treatments

with a total of 20 sessions (10 sessions per treatment), with each session lasting 30 periods. Ten subjects participated in each session. Subjects were paid for three randomly selected periods, and the average individual earnings were £14. All instructions can be found in Appendix A.

The experimental design consists of two treatments: the No-Moving Treatment (NMT) and the Moving Treatment (MT). Treatments differ depending on whether the groups are endogenously formed.

No Moving Treatment (NMT). In the first period, subjects are randomly assigned to one of two groups – A and B – of 5 people each. Each period is composed of two (sequential) stages: a *contribution stage* and a *punishment stage*. In the *contribution stage*, each subject decides how to allocate the 50 tokens she is endowed with between a group account and a private account.

There are two institutional settings (rule-sets) affecting subjects' payoffs.

(a) Under **IS**, informal sanctioning institutions are in place. In the *punishment stage*, subjects observe the contributions to the group account made by their fellow group members. They then have the opportunity to reduce the earnings of the other members of their group. Specifically, any subject i can impose a sanction of 3 units on any other group member j at a cost of 1 unit to herself. A subject can spend at most 10 tokens on punishing any other. The payoff of subject i under IS is:

$$\pi_i = (50 - C_i) + \frac{1.6}{n^h}\left(\sum_{j=1}^{n^h} C_j\right) - \sum_{j=1}^{n^h} R_{i,j} - 3\sum_{j=1}^{n^h} R_{j,i}, \tag{1}$$

for $i, j = 1, ..., n^h$ and $h = A, B$, where n^h denotes the total number of subjects located in group h, and $C_i \in [0, 50]$ denotes i's contribution to the group account. We denote by $R_{i,j}$ ($R_{j,i}$) the number of tokens that subject i (j) uses to punish subject j (i).[3]

[3]When $n^h = 1$, the lone subject in a group simply receives his or her endowment of 50 tokens.

Finally, the factor multiplying the contributions to the group account is equal to 1.6, implying a marginal per-capita return from the group-h account equal to $\frac{1.6}{n^h}$.

(b) Under **FS**, formal sanctions are in place. Each subject in the group pays a fixed fee of 5 tokens per period. In addition, in the *punishment stage*, each subject pays a fine equal to 80% of the amount of tokens allocated to the private account in a given period. The fixed fee and the fine (if applicable) are deducted from subjects' monetary payoff. Therefore, under FS, subject i's per-period monetary payoff is:

$$\pi_i = (50 - C_i)(1 - 0.8) + \frac{1.6}{n^h}\left(\sum_{j=1}^{n^h} C_j\right) - 5, \tag{2}$$

for $i, j = 1, ..., n^h$ and $h = A, B$.[4]

At the end of the *punishment stage*, each subject learns her/his payoff. In IS, subjects also learn the amount of punishment they receive, but not who punished them or how much punishment others receive in total. In FS, subjects are informed of the amount of the fine they pay. In both institutional settings, participants also receive information about: *(i)* the average contribution to the group account in their current group; *(ii)* the average payoffs in their current group; *(iii)* the average payoffs in the other group; *(iv)* the rule-set implemented in their current group; and *(v)* the rule-set implemented in the other group.

In the first 5 periods, both Group A and B use IS. Starting from period 6, subjects vote every 5 periods on the rule-set to be implemented in their group. Subjects therefore vote 5 times in total, with the *voting stage* taking place at the beginning of the period. The rule-set is implemented immediately after voting and applies until the next voting round. Hence, from period 6 onward, subjects first decide on the rule-set (if the period includes a *voting stage*), then contribute to the group account (*contribution stage*), and finally participate in the *punishment stage*.

[4]Note that in the experimental instructions (Appendix A) we refer to the IS (respectively, FS) rule-sets as Rule Set 1 (respectively, Rule Set 2).

Moving Treatment (MT). This treatment differs from NMT in one aspect only. In NMT, subjects cannot move between groups: there are two groups of 5 members that are fixed for the entire duration of the experiment. In MT, by contrast, Group A and B are fixed in the first 5 periods, but starting from period 6 moving between groups is allowed at the end of each period. Subjects enter the *moving stage* after the *punishment stage*, and simultaneously decide whether to move from their current group. Hence, from period 6 onwards, subjects first decide on the rule-set to be implement in their current group (if the period includes a *voting stage*), then contribute to the group account (*contribution stage*), then participate in the *punishment stage*, and finally decide whether to move.[5]

As explained in the Introduction, the decision to keep groups fixed and let all groups use IS during the first 5 periods of the experiment was taken in order to allow groups to develop distinct, cooperative norms in the IS environment. This is necessary in order to test the conjecture that reactions to opening up the possibility of migration depend on pre-existing norms of cooperation.

3 Results

We first compare subjects' voting behavior in our baseline and main treatments, and then examine the main determinants of subjects' contribution, punishment, and migration decisions.

3.1 Voting Behavior

We start with an overview of group dynamics in voting behavior in the two treatments. Figure 1 plots the average percentage of people voting for IS over time in both MT and NMT. The figure reveals that a larger fraction of the population chooses IS in all voting periods under NMT than under MT. As a result, the average percentage of subjects voting for IS

[5]In MT, subjects are informed about their current group size at the beginning of each period.

across all periods is significantly larger in NMT (57.2%) than in MT (46.4%).[6]

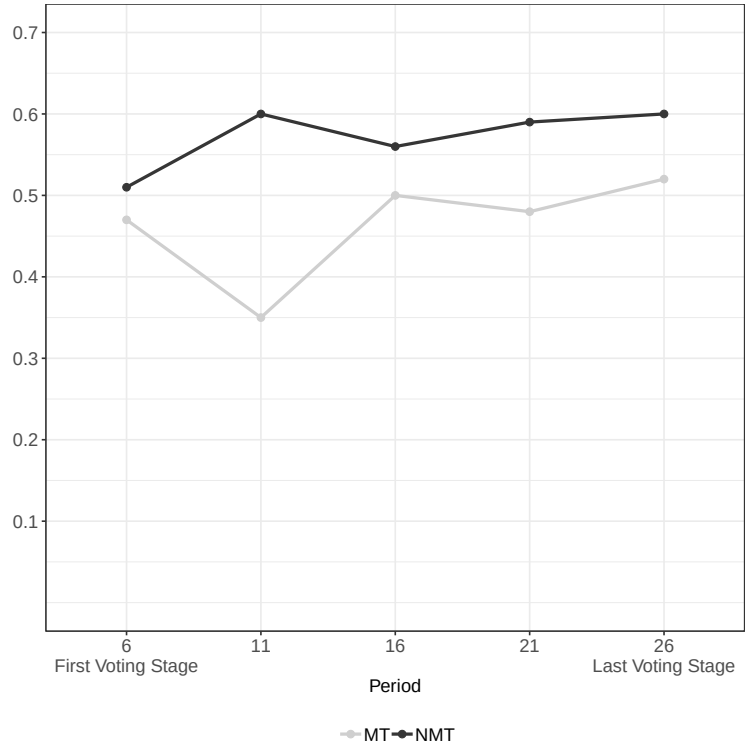

Figure 1: Proportion of participants voting for IS over time, by treatment. Gray (black) lines represent the proportion of subjects voting for IS under MT (NMT) in each of the periods in which a voting stage takes place, averaged across sessions.

To study whether voting depends on initial levels of cooperation in the IS regime, Figure 2 divides groups depending on whether their average contribution to the public good in the first five periods of each session is "high" or "low". In particular, the average contribution is defined as "high" (resp., "low") if it is above (resp., below) the median sample contribution – i.e., the median contribution of all the groups in the experiment – in these initial periods, when institutions and group affiliations are fixed. For both treatments, the top panel of the figure plots the average share of subjects originally located in groups with "high" initial contributions who vote for IS. The bottom panel shows the average share of subjects in groups

[6] $\chi^2 = 11.25$, $p = 0.00$, for a two-sided test for equality of proportions.

originally located with "low" initial contributions voting for IS under the two treatments.

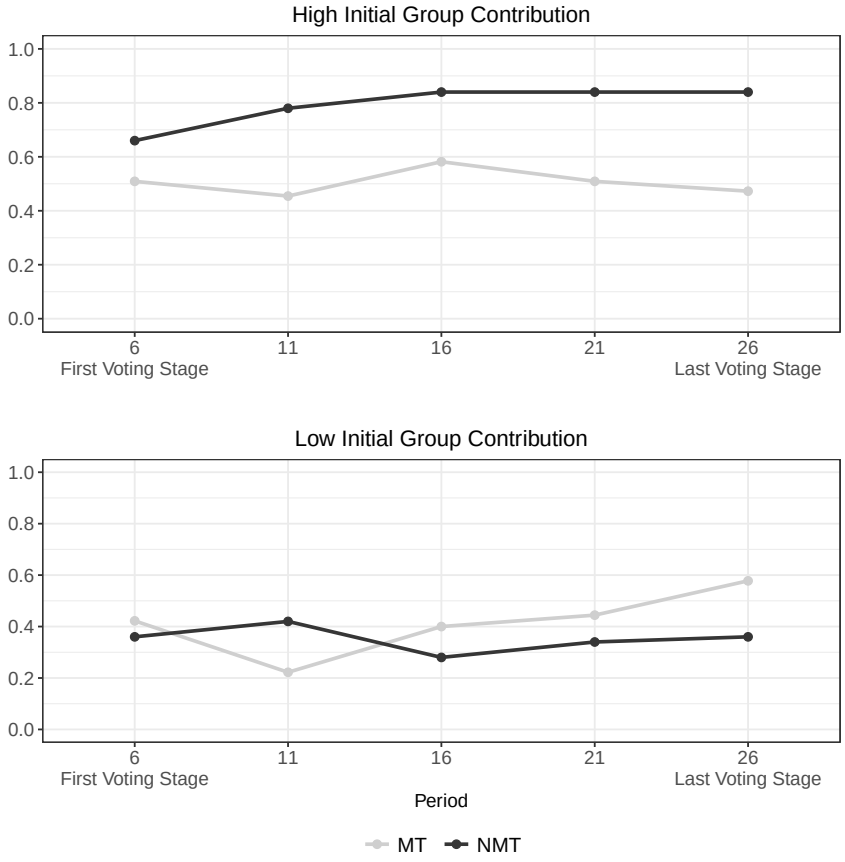

Figure 2: Proportion of participants voting for IS over time, by group's average initial contribution and treatment. Gray (black) lines represent the proportion of subjects voting for IS under MT (NMT) in each of the periods in which a voting stage takes place, averaged across sessions.

The top panel reveals that participants whose group exhibited relatively high initial contribution levels are more likely to vote for informal sanctioning institutions when the groups are closed than when they are open. The percentage of subjects who vote for IS in the last three voting stages is significantly larger in NMT (84%) than in MT (54%).[7,8] The bottom panel of the figure shows that, for participants whose group exhibited relatively

[7] $\chi^2 = 31.56$, $p = 0.00$ for a two-tailed test of equality of proportions taking each session as an independent observation. Equivalently, Mann-Whitney test: $z = 2.33$, $p = 0.02$, two-tailed.

[8] We choose the last three periods that include a voting stage to allow for some learning in the process. Results are the same if we take the last two periods that include a voting stage ($\chi^2 = 23.34$, $p = 0.00$, two-tailed test of equality of proportions; Mann-Whitney test: $z = 2.28$, $p = 0.02$, two-tailed) and the last voting stage ($\chi^2 = 11.57$, $p = 0.00$; Mann-Whitney test: $z = 2.27$, $p = 0.02$).

low initial contributions, the results reverse: subjects are less likely to vote for IS when the groups remain the same over the course of the experiment. Considering again only the last three voting periods, the corresponding proportions for NMT and MT are 32.7% and 46.0%, respectively.[9] Similar patterns hold using alternative definitions of groups with (relatively) "high" initial contributions.[10]

Result 1. *Over the course of the experiment, a larger proportion of subjects who were originally located in a group with "high" initial levels of cooperation vote for informal sanctioning institutions (IS) in closed societies (NMT) than in open ones (MT). The opposite result holds for subjects who were originally located in a group with "low" initial levels of cooperation: in this case, a larger share of subjects vote for informal sanctioning institutions (IS) in open societies (MT) than in closed ones (NMT).*

To better understand the mechanisms underlying the group-level patterns reported above, we next present individual-level analyses of voting decisions. The first column of Table 1 reports marginal effects (in percentage points) computed from a panel probit model fitted to data from the five periods that include a voting stage.[11] The dependent variable, $Vote_{i,t}$, equals 1 (0) if subject i voted for (against) IS in period t. The explanatory variables include: $Moving_i$, a binary covariate that equals 1 if i participated in MT, and 0 otherwise; $Contribution_{i,t-1} - Contribution_{g(-i),t-1}$, the difference between i's contribution and the average contribution of the other members of her group in period $t-1$; $IS_{i,t-1}$, a

[9] $\chi^2 = 5.04$, $p = 0.02$ for a two-sided test for equality of proportions.

[10] To assess the robustness of our findings, we replicated this analysis - as well as all those discussed below - identifying a group as having a "high" initial contribution when the average contributions of its members in the first five periods: (i) are greater than a pre-specified value (e.g, 30, 35 tokens); or (ii) are above the third quartile of the distribution of initial contributions across all groups in the experiment; or (iii) exceed the average initial contributions of all the groups. The main findings continue to hold regardless of the particular measure used (see Figure B.1 in Appendix B).

[11] We report the "raw" parameter estimates in Table B.1 of Appendix B. These must be transformed to obtain estimates for the marginal effects – i.e., the change in predicted probability associated with changes in the explanatory variables (Greene, 2003, p. 667).

dichotomous variable taking the value 1 if i played under the informal sanctioning mechanism in the period preceding each voting stage, and 0 otherwise; and *Group Size*$_{g(i),t}$, the size of i's group in period t. We also incorporate covariates capturing the behavior of the members of the two groups in the first five periods of the experiment: *High Initial Group Contribution*$_{g(i)}$, an indicator taking the value 1 when the average contribution in i's group in those initial periods was greater than the median initial contribution of all the groups in the experiment, and 0 otherwise; the interaction between *High Initial Group Contribution*$_{g(i)}$ and $Moving_i$; *High Initial Group Payoff*$_{h \neq g}$, a dummy indicating whether the average payoff of the members of the other group - i.e., the group subject i did not originally belong to - in the first 5 periods exceeded the median payoff of all the groups in those initial periods; and the interaction between *High Initial Group Payoff*$_{h \neq g}$ and $Moving_i$.[12] In addition, the model includes subject-specific (correlated) random effects (Wooldridge 2005, 2010) to account for time-invariant individual heterogeneity, as well as period, group and session random intercepts controlling for temporal shocks affecting all subjects, for contemporaneous correlation between same-group members, and for potential session effects, respectively (Poen, 2009; Fréchette, 2012).[13]

The estimates for *High Initial Group Contribution*$_{g(i)}$ imply that, in NMT, the tendency to vote for IS increases significantly when subjects experience high initial group contributions. On average, the probability of voting for IS is about 37 percentage points higher for individuals who belonged to groups that made relatively high contributions at the beginning of the experiment than for subjects who were part of groups that contributed relatively little in the first five periods. Note that informal sanctions are in place in both groups in these

[12] As mentioned in Section 2, subjects receive information about the average contribution to the public group in their own group, but not in the other group. They are informed about the average payoffs in both groups, though. For completeness, we replicated the analyses replacing *High Initial Group Contribution*$_{g(i)}$ with *High Initial Group Payoff*$_{g(i)}$; the main findings remain unchanged (see Table B.2 in the Appendix).

[13] As is well known, (bias-corrected) fixed effects estimators for panel probit models (e.g., Fernández-Val and Weidner, 2016) cannot identify the coefficients of time-invariant covariates. For robustness, we also applied the estimator proposed by Kripfganz and Schwarz (2019) for fixed-effects linear panel models with time-invariant regressors. The results are aligned with those from the random effects probit models (see Table B.3 in the Appendix).

Table 1
Determinants of individual voting decisions

	(1)	(2)	(3)
$Moving_i$	12.07	11.83	9.71
	(9.67)	(8.48)	(7.68)
$High\ Initial\ Group\ Contribution_{g(i)}$	36.62***	29.45***	25.14***
	(6.55)	(6.95)	(6.63)
$High\ Initial\ Group\ Contribution_{g(i)} \times Moving_i$	-31.24***	-26.13***	-21.64**
	(10.96)	(9.88)	(9.08)
$High\ Initial\ Group\ Payoff_{h \neq g}$	3.73	5.40	4.41
	(6.61)	(5.57)	(5.20)
$High\ Initial\ Group\ Payoff_{h \neq g} \times Moving_i$	-10.03	-8.86	-7.22
	(9.05)	(8.12)	(7.88)
$Contribution_{i,t-1} - Contribution_{g(-i),t-1}$	-0.57**	-0.36	-0.18
	(0.24)	(0.25)	(0.12)
$IS_{i,t-1}$	16.87***	9.77**	11.60***
	(4.19)	(4.23)	(4.39)
$Group\ Size_{g(i),t}$	-1.37	-0.85	-1.07
	(0.97)	(0.88)	(0.87)
$Vote_{i,t-1}$		28.39***	28.83***
		(4.80)	(4.74)
$Payoff_{i,t-1}$			-0.49**
			(0.19)
$Payoff_{i,t-1} \times IS_{i,t-1}$			2.39***
			(0.48)
Observations	1,000	800	800
Log likelihood	-549.47	-393.09	-378.89

Notes. The table reports the change in the probability of voting for informal institutions (in percentage points) associated with a change in the covariates; units of observation are individuals-per-period (with the sample restricted to the five periods that include a voting stage). All specifications include subject, period, group, and session random effects. Standard errors are presented in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

first five periods. Thus, this result suggests that participants are more willing to vote for IS when these institutions worked properly in the first place.

The negative and significant effect of the interaction between $High\ Contribution_{g(i)}$ and $Moving_i$ shows that, in line with the results reported in Figure 2 (see *Result 1*), participants originally located in groups with relatively high initial contributions are more likely to vote for IS when they are in a society with exogenously fixed groups than when they are in societies

with endogenous group formation. In fact, the positive relationship between *High Initial Group Contribution*$_{g(i)}$ and $Pr(Vote_{i,t} = 1)$ is exclusively driven by the voting behavior of subjects in NMT. In groups with low initial contributions, the mere possibility of moving between groups does not have a systematic influence on individuals' voting behavior. This is also observable from the statistically insignificant main effect of $Moving_i$ in the first row of Table 1.

The probability of voting for informal sanctioning institutions also increases for subjects experiencing these institutions right before any given voting stage: the estimated marginal effect for $IS_{i,t-1}$ implies that subjects playing under informal institutions are almost 17 percentage points more likely to vote for IS in the following period than those who experienced formal sanctions. The probability of voting for IS also decreases with the difference between the subject's contribution and the contribution of the other members of her group in the period immediately preceding a vote: each additional token a subject contributes above her group's average is associated with a 0.57 percentage point decrease in $Pr(Vote_{i,t} = 1)$. In other words, people become more (less) willing to adopt formal (informal) sanctioning institutions when their contribution is high compared to the average contribution of their group.

The estimates in column (1) also indicate that the probability that a subject i votes for IS is not systematically influenced by the average payoff of the members of the other group in the first 5 periods under either treatment (MT or NMT).[14] Finally, the fact that the marginal effect of *Group Size*$_{g(i),t}$ on $Pr(Vote_{i,t} = 1)$ is statistically indistinguishable from zero is consistent with the results of Gürerk, Irlenbusch and Rockenbach (2006): an increase in the size of the group does not reduce the probability that participants support the implementation of informal sanctions.

To account for autocorrelation and for subjects' tendency to adjust their behavior over

[14] We repeated the analysis replacing the *High Initial Group Payoff*$_{h \neq g}$ dummy with a continuous variable measuring the average payoff (in tokens) of the members of the other group in the first 5 periods of the experiment. The results remain unchanged, i.e. neither the average payoff of the members of group $h \neq g$ nor its interaction with $Moving_i$ has a statistically significant marginal effect on $Pr(Vote_{i,t} = 1)$.

time (Smith, 2013), column (2) in Table 1 adds the lag of the dependent variable to the regressors of the column (1) specification. The estimate for $Vote_{i,t-1}$ indicates that, holding everything else constant, the average propensity to support IS is more than 28 percentage points higher for subjects who already voted in favor of that type of sanctions in the previous voting stage. This result suggests that subjects are consistent in their preferences over time. The estimates for most of the other explanatory variables are in line with those reported in column (1), the only exception being $Contribution_{i,t-1} - Contribution_{g(-i),t-1}$, which ceases to have a statistically significant effect on $Pr(Vote_{i,t} = 1)$ in this specification.

Column (3) adds $Payoff_{i,t-1}$, subject i's payoff in the period previous to voting, along with its interaction with $IS_{i,t-1}$, to the set of explanatory variables of the column (2) specification. An increase of one token in the average subject's payoff in the period preceding a voting stage is associated with a 0.49 percentage point decrease in her willingness to vote for informal sanctions if such payoffs were generated under FS, and with a 2.39 percentage point increase if the earnings were generated under IS.

To further study whether members of groups with high initial levels of cooperation worry about inward migration of low contributors and therefore tend to vote for FS when societies are open, we now focus specifically on the first voting stage. Moreover, we consider only "heterogeneous sessions", i.e. sessions where one group had high relative contributions in the initial 5 periods while the contributions of the other group were below the median sample value. We would expect the interaction between openness and initial average contributions to be particularly strong in these sessions, since high-contribution groups in open societies face a high likelihood of inward migration by members of low-contribution groups. Additionally, initial contribution levels are expected to be most salient in the first period that includes a voting stage.

Figure 3 plots the percentage of subjects voting for IS in the first voting round of each treatment in these "heterogeneous sessions", distinguishing, as in Figure 2, between groups with (relatively) "high" and "low" initial contributions. Since subjects are not allowed to

move between groups before the first voting stage under either treatment, when migration is allowed (in MT), subjects' voting behavior is just affected by their expectation of out-members' future behavior (but not by the direct effect on i's payoffs of out-members' past contributions *per se*).

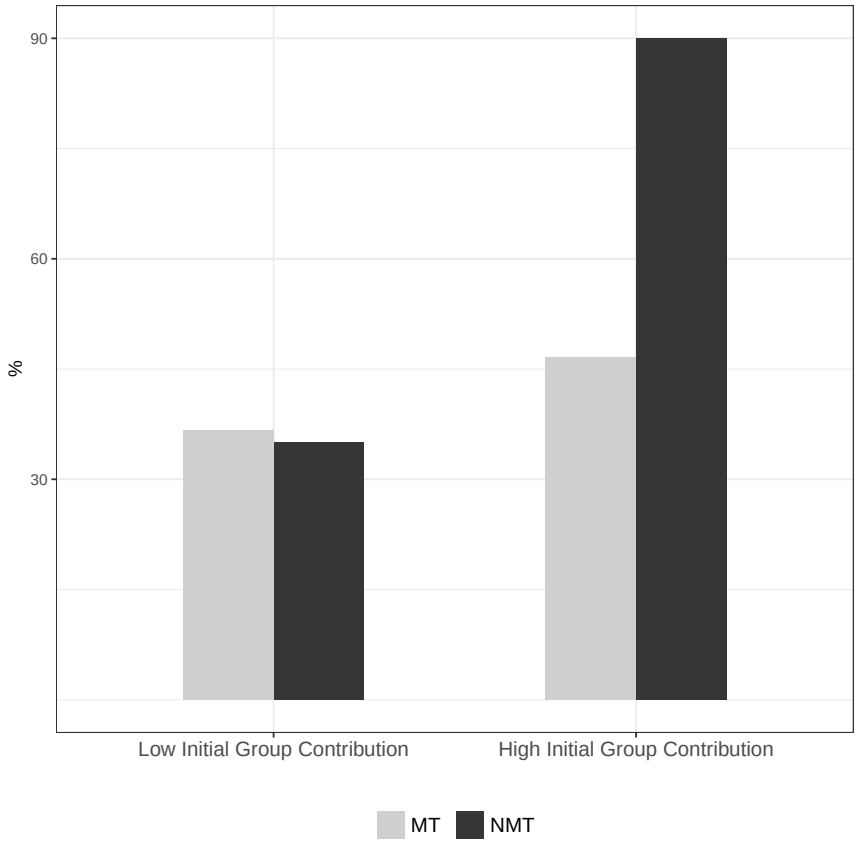

Figure 3: First vote on IS in "heterogeneous sessions". The figure plots the proportion of subjects voting for IS in "high" and "low" contribution groups in the first voting stage under both MT (gray bars) and NMT (black bars). The analysis is restricted to sessions in which one group exhibits (relatively) high initial contribution while the average initial contribution of the other group is below the sample median.

The figure shows that the voting behavior in the two treatments is very similar for groups with "low" initial contribution levels: the percentage of subjects voting for IS in MT is 36.7%, compared to a 35% in NMT. The two proportions are statistically indistinguishable

($\chi^2 = 0.01$, $p = 0.90$, for a two-sided test for equality of proportions). Results however differ for groups with "high" initial contribution. In this case, 90% of subjects vote for IS when groups are fixed (NMT), whereas less than 47% vote for IS in MT – a statistically significant difference ($\chi^2 = 9.11$, $p = 0.00$, two-sided). As expected, this difference is even larger than the corresponding effects reported in Figure 2 for "high" contributors. These results support the view that subjects' expectations of other participants' behavior play a significant role in institutional choice.

3.2 Contributions and Punishment

Next, we analyze subjects' contribution to the public good and their punishment decisions in IS.[15] This analysis is key to a better understanding of subjects' propensity to vote for different sanctioning institutions.

Figure 4 plots the evolution of the average individual contribution and the average punishment received by subjects when groups implement informal (upper panel) and formal (lower panel) sanctioning institutions under the two treatments.[16] As seen in the upper panel of the figure, when informal sanctioning institutions are in place, contributions are higher in NMT than in MT. In NMT, the average individual contribution across all sessions and periods is 44.69 tokens, whereas the corresponding value in MT is 37.46. This difference between treatments is statistically significant (Mann-Whitney test: $z = 3.89$, $p = 0.00$, two-tailed).[17] Moreover, under IS, we observe a significant monotonic, positive trend in contributions in both MT and NMT (Mann-Kendall trend test: z=3.62, p=0.00 and z=3.46, p=0.00 for MN and NMT, respectively; both two-sided tests). Although the

[15] For completeness, we also examine subjects' earnings, which are strongly related to their contributions and punishment decisions. These results are summarized in Appendix B (Figure B.2 and Table B.4).

[16] Obviously, the fact that rule-sets are endogenous complicates the attribution of a casual relationship between the institutional setting and subjects' contribution and punishment decisions. We can nonetheless assess whether there is a systematic association between the type of institutions in place and the average contribution/punishment levels.

[17] This comparison considers only the average contributions from period 6 onwards. By contrast, average contributions are statistically indistinguishable between treatments in the first 5 periods of the experiment, when institutions and group affiliations are fixed (Mann-Whitney test: $z = 1.28$, $p = 0.20$, two-tailed).

average contribution across sessions is also somewhat higher in NMT (45.12) than in MT (42.8) under FS, the difference between treatments is statistically indistinguishable from zero (Mann-Whitney test: $z = 0.65$, $p = 0.52$, two-tailed). In this case, we observe that only contributions under NMT exhibit a significant monotonic trend (Mann-Kendall trend test: z=0.68, p=0.50, and z=3.76, p=0.00 for MN and NMT, respectively).

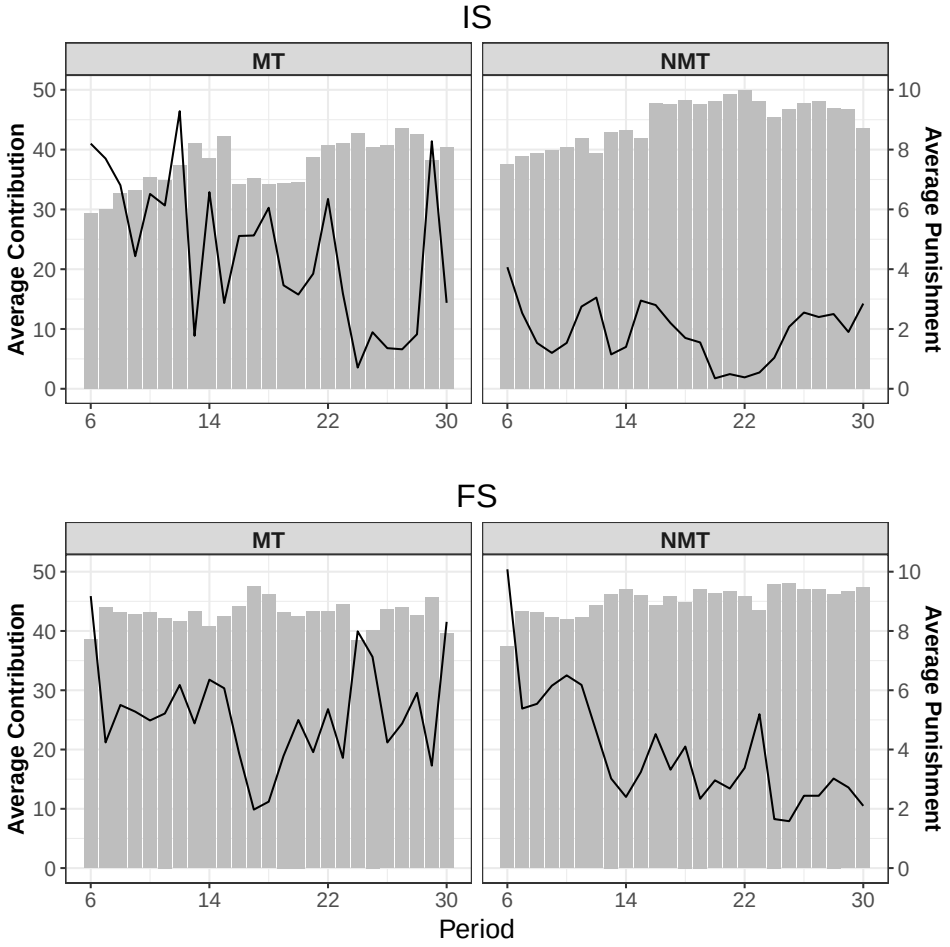

Figure 4: Average individual contribution and punishment received over time, by rule-set and treatment. The upper (lower) panel plot the average contribution and punishment received by individuals located in groups implementing IS (FS) under the two treatments, in tokens. Bars represent the average contribution per period (measured on the left axis), while solid lines correspond to the average punishment received (measured on the right axis).

Altogether, Figure 4 shows that the overall contribution levels are significantly larger under FS (43.60) than under IS (38.10) (Wilcoxon signed rank test: $z = 2.45$, $p = 0.01$, two-tailed).

Result 2. *When informal sanctioning institutions are in place, contributions are higher in closed societies than in open ones. By contrast, when formal sanctioning institutions are adopted, societies are able to support high contributions to the public good in both treatments. Overall, contributions under formal sanctioning institutions are significantly larger than under informal ones.*

To study the main determinants of contributions, column (1) of Table 2 reports the marginal effects from a (doubly) censored regression model (Greene, 2003) in which the dependent variable is i's individual contribution in period t, $Contribution_{i,t}$.[18] The explanatory variables include some of the predictors used in the individual-level analyses of Table 1 – $Moving_i$, $High\ Initial\ Group\ Contribution_{g(i)}$, $Contribution_{g(-i),t-1}$, $Group\ Size_{g(i),t}$, IS_t – along with the amount of punishment received by subject i in the preceding period, a dummy indicating whether i played under informal sanctions in period $t-1$, the interaction between $Punishment\ Received_{i,t-1}$ and $IS_{i,t-1}$, and the lag of the dependent variable (in order to control for persistence in individuals' contribution decisions). The model also incorporates subject, period, group and session random effects.[19]

Individual contributions to the public account increase significantly for subjects who originally belonged to a group with relatively "high" initial contributions: each additional token contributed by i's group in the first five periods of a session is associated with an increase of 1.29 tokens in the average subject's contribution at a later period t. All else

[18] "Raw" parameter estimates - representing the effect of the predictors on the uncensored latent variables - are reported in Appendix B (Table B.5).

[19] The results are generally similar using the two-step method proposed by Honoré and Kesina (2017) to estimate fixed-effects censored regression models with time-invariant explanatory variables (Table B.6 in Appendix B).

Table 2
Determinants of individual contributions

	(1)	(2)
$Moving_i$	-0.32	-0.76
	(0.70)	(0.67)
$IS_{i,t}$	-3.41***	-3.33***
	(0.45)	(0.45)
$Moving_i \times IS_{i,t}$	-1.27**	-1.22**
	(0.51)	(0.50)
$High\ Initial\ Group\ Contribution_{g(i)}$	1.29**	1.01*
	(0.55)	(0.54)
$Contribution_{g(-i),t-1}$	0.21***	0.24***
	(0.02)	(0.02)
$Group\ Size_{g(i),t}$	0.34***	0.16*
	(0.09)	(0.09)
$Contribution_{i,t-1}$	0.10***	0.11***
	(0.02)	(0.02)
$Punishment\ Received_{i,t-1}$	-0.03	-0.01
	(0.03)	(0.02)
$IS_{i,t-1}$	1.81***	2.22***
	(0.39)	(0.41)
$Punishment\ Received_{i,t-1} \times IS_{i,t-1}$	0.05**	
	(0.02)	
$Punishment\ Received_{i,t-1} \times \mathrm{IS}_{i,t-1} \times$		0.12
$\mathbb{1}\big(Contribution_{i,t-1} > Contribution_{g(-i),t-1}\big)$		(0.08)
$Punishment\ Received_{i,t-1} \times \mathrm{IS}_{i,t-1} \times$		0.04**
$\mathbb{1}\big(Contribution_{i,t-1} < Contribution_{g(-i),t-1}\big)$		(0.02)
Observations	4,800	4,800
Log likelihood	-11,558.33	-11,368.93

Notes. The table reports marginal effects for the covariates included in the panel doubly-censored regression models for individual contributions; units of observation are individuals-per-period. Both specifications include subject, period, group, and session random effects. Standard errors are presented in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

equal, i's contributions also increase with the contribution of the other members of her group in the previous period: the estimated marginal effects of $Contribution_{g(-i),t-1}$ imply that each additional token contributed by other subjects located in i's group in the previous period leads to a subsequent increase in her own contribution of 0.21 tokens. These findings

imply that subjects' contribution behavior depends on their previous experience: individuals located – either at the beginning of the experiment or in a more recent period – in a group that cooperates more tend to contribute themselves more to the public good later on. The estimate for $Contribution_{i,t-1}$ reveals also a moderate but significant degree of persistence in individuals' propensity to cooperate: roughly 10% of subject i's contribution in period t is explained by her contribution in the previous period. Group size has a positive and significant effect on $Contribution_{i,t}$ as well.

On the contrary, in NMT i's contribution declines by 3.41 tokens on average when her group is governed by IS, in line with *Result 2*. Furthermore, and also consistent with *Result 2*, the drop in contributions under informal sanctioning institutions is even more marked in open societies, as indicated by the negative and significant interaction between $Moving_i$ and $IS_{i,t}$.

Concerning the effect of punishment, the marginal effect of *Punishment Received*$_{i,t-1}$ on i's contribution in period t is statistically indistinguishable from zero in FS. By contrast, the estimate for the interaction between *Punishment Received*$_{i,t-1}$ and $\text{IS}_{i,t-1}$ is positive and significant, indicating that punishment is on average more effective in raising contributions in IS than in FS.[20] To account for the potential impact of "perverse" punishment (i.e., punishment of high contributors) on contributions, column (2) re-estimates the relationship between *Punishment Received*$_{i,t-1}$ and $Contribution_{i,t}$ in IS, distinguishing between subjects who contributed more or less than their group mean.[21] The estimate for *Punishment Received*$_{i,t-1} \times \text{IS}_{i,t-1} \times \mathbb{1}\big(Contribution_{i,t-1} < Contribution_{g(-i),t-1}\big)$ reveals a positive and significant relationship between the amount of informal punishment received in $t-1$ and subsequent contributions for individuals whose allocation to the group account exceeded the group average. By contrast, greater monetary sanctions do not have a systematic impact on

[20] It is worth noting that, as shown in Figure 4, average contributions are quite high in FS (and systematically higher than in IS). Hence, there is relatively "little room" for further increases. This may contribute to explain the insignificant marginal effect of *Punishment Received*$_{i,t-1}$ on $Contribution_{i,t}$ in FS.

[21] Perverse punishment may be present in IS (cf. e.g. Herrmann, Thöni and Gächter, 2008; Ertan, Page and Putterman, 2009), but is ruled out – by definition – in FS.

$Contribution_{i,t}$ when i contributed more than her fellow group members.

We end this section by investigating punishment behavior. We focus specifically on punishment under IS, since the relationship between contribution and punishment is straightforward in FS (see Section 2). The top panel of Figure 4 revealed that the average punishment received by subjects in groups implementing informal sanctioning institutions is lower in closed than in open societies – mainly because contributions are larger in NMT than in MT. The results presented in column (1) of Table 3, which reports marginal effects estimated from a random effects censored regression model examining the determinants of the amount of punishment received by subject i in period t, are in line with this finding.[22] Holding individual contributions fixed, the punishment received by an average subject is systematically larger in MT than in NMT.

Table 3
Determinants of the amount of punishment received under IS

	(1)	(2)
$Moving_i$	1.20**	0.34
	(0.50)	(0.51)
$Contribution_{i,t}$	-0.19***	-0.16***
	(0.01)	(0.02)
$Contribution_{i,t} \times Moving_i$		0.04**
		(0.02)
$max(Contribution_{g(-i),t}$ - $Contribution_{i,t}, 0)$		0.09***
		(0.02)
$max(Contribution_{i,t}$ - $Contribution_{g(-i),t}, 0)$		0.01
		(0.01)
Observations	2,399	2,399
Log likelihood	-4,573.076	-4,394.99

Notes. The table reports marginal effects for the covariates included in panel doubly-censored regression models examining the determinants of the amount of punishment received by subject i in period t under IS. Units of observation are individuals-per-period. Both specifications include subject, period, group, and session random effects. Standard errors are presented in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Column (2) expands the column (1) specification by including an interaction between $Moving_i$ and i's contribution in t, along with the negative and positive deviation of i's

[22] "Raw" parameter estimates - representing the effect of the predictors on the uncensored latent variables - are reported in Appendix B (Table B.7). The substantive results are similar using Honoré and Kesina (2017)'s fixed-effects estimator (Table B.8).

contribution from the average contribution of the other members of her group (measured in absolute values). Although the marginal effect of $Moving_i$ is no longer significant, the estimate for $Contribution_{i,t} \times Moving_i$ indicates that, for any given contribution level, the amount of punishment received by the average subject is consistently higher in MT than in NMT. This provides some evidence that, in open societies, subjects may feel the need to be somewhat harsher in their punishment behavior, perhaps as a way to establish ground rules and discipline newcomers.

As for the impact of deviations from other group members' contributions, the positive and significant estimate for for $max(Contribution_{g(-i),t} - Contribution_{i,t}, 0)$ indicates that the punishment received by i increases the further below the group average is her contribution. In contrast, the size of the positive deviation has no systematic impact on i' punishment, as the marginal effect of $max(Contribution_{i,t} - Contribution_{g(-i),t}, 0)$ is statistically indistinguishable from zero. These findings underscore that there is more pro-social than anti-social punishment in IS, suggesting that informal punishment is mostly well-targeted at free riders.[23]

3.3 Migration Behavior

We now analyze the main determinants of subjects' location choice in MT. This analysis sheds light on the endogenous dynamics of group formation. It also informs our results regarding the lack of significance of group size on subjects' voting behavior (Section 3.1). Groups size can vary more or less abruptly during the course of the experiment. Arguably, a "slow" rate of growth makes it easier for a host society implementing IS to discipline newcomers.[24] Hence, we expect that a "slow" increment in the number of subjects populating a given group will minimize any potentially negative impact of group size on the effectiveness of informal sanctioning institutions to induce high contributions. This, in fact, could explain

[23] This is also observed in Figure B.3 in Appendix B, which displays the punishment directed towards subjects in IS who contribute more than the group average, as a share of the total punishment in each period.

[24] Evidence pointing in this direction is provided by Weber (2006). The author finds that slow growth in group size has a positive effect on coordination.

why people do not vote for IS significantly less when the group size increases.

The top left panel of Figure 5 plots the average share of subjects moving between groups in each period. A strictly positive fraction of subjects migrate in any given period and 20% on average switch groups each period over the course of the experiment, when moving is allowed. The top right panel, in turn, distinguishes between moves towards and moves away from IS when the two groups are implementing different rule-sets at the time those moves take place. Considering only periods/sessions in which one group implements IS while the other group has FS in place (ca. 42% of the observations), we find that 79% of the movements go from FS to IS, while only 21% of the group switches go in the opposite direction. These differences are statistically significant (χ^2=64.98, p=0.00 for a two-sided test for equality of proportions).[25]

The bottom panels of the figure, which plot the average number of members in each group during the course of the experiment, indicate that subjects' migration patterns lead to groups diverging in size over time. The average group size across all sessions and over the entire duration of the experiment is 7.03 for Group A and 2.97 for Group B, where, for ease of exposition, we label the group with the larger (smaller) number of members in each period of each session as A (B). In both cases, the average group size is significantly different from 5 (Mann-Whitney test: $z = 4.03$, $p = 0.00$, two-tailed). The maximum group size in a given period (averaged across sessions) is 8.7. The mean per-period change in the number of members of a group – averaged across sessions – is 0.4. Hence, considering that moving is entirely costless, group size arguably grows (or shrinks) rather slowly in our experiment, which might help explain the insignificant marginal effect of $Group\ Size_{g(i),t}$ on $Pr(Vote_{i,t} = 1)$ reported in Table 1.

To analyze subjects' migration choices in greater detail, Table 4 reports marginal effects

[25] A possible explanation for this result is that IS is implemented because it functions relatively well. If this is the case, since IS is a relatively efficient institution with respect to FS, it tends to attract more subjects. Figure B.4 in the Appendix complements this analysis by plotting the share of participants located in a group with IS in each period. This would be the consequence of the combination of voting and migration decisions.

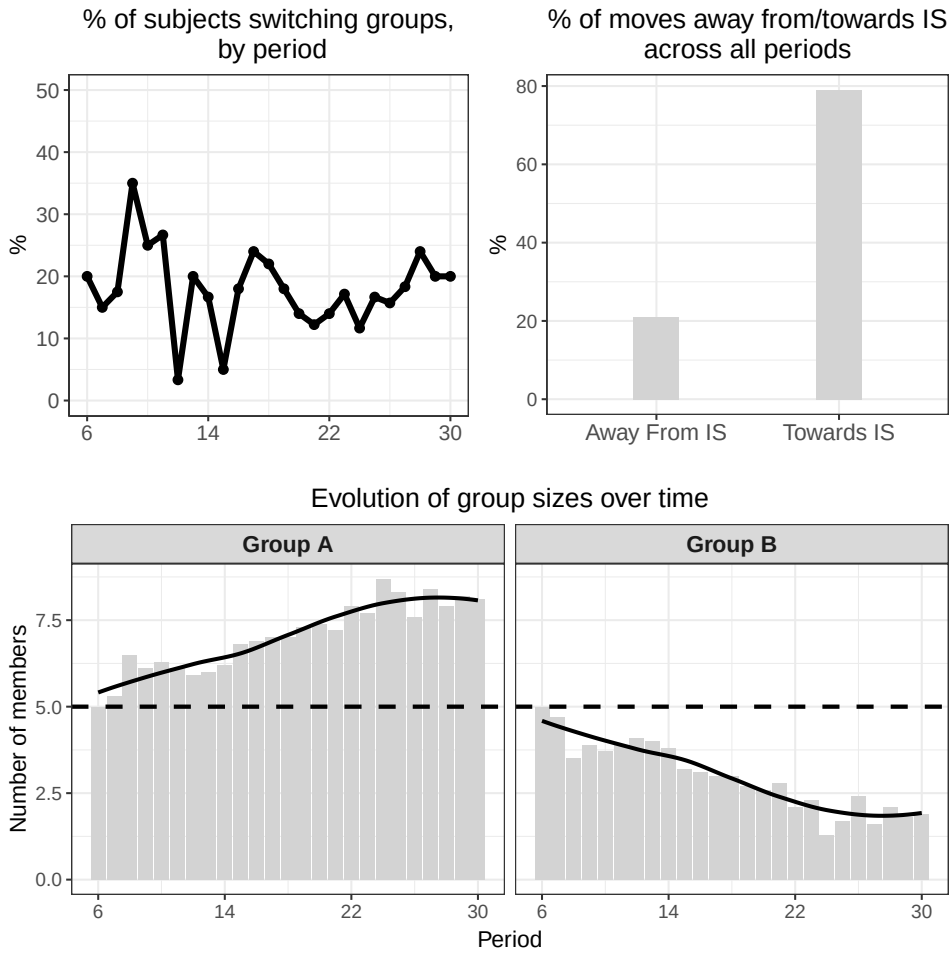

Figure 5: Migration patterns and group sizes over time. The upper-left panel of the figure plots the average proportion of subjects moving between groups in each period (averaged across all sessions). The upper-right panel plots the proportion of moves away from/towards groups that are implementing sanctions in the period in which the moves take place – out of the total number of movements between groups with different rule-sets in that period, across all sessions. Finally, the bottom panel shows the number of members of Group A (the larger group) and Group B (the smaller group) by period, averaged across sessions.

from a random effects panel probit model for migration decisions, in percentage points.[26] The dependent variable is $Migration_{i,t}$, a dummy taking the value 1 if subject i moves in period t, and 0 otherwise. The predictors in column (1) are essentially the same as those used

[26]Parameter estimates are reported in Appendix B (Table B.9). Since these specifications do not include time-invariant predictors, we also fitted two-way (subject and period) fixed-effects probit models using the method developed by Fernández-Val and Weidner (2016). The main results remain qualitatively similar (see Table B.10).

in previous individual-level analysis, with the addition of: *Average Payoff* $_{h\neq g,t}$ − *Average Payoff* $_{g,t}$, the difference between the average payoff of the members of the other group and the mean earnings of the subjects located in i's group in period t; and *Vote Different from Group* $_{i,t-1}$, an indicator measuring whether i's institutional choice in the most recent voting period differed from the decision of the majority of her group members.

Table 4
Determinants of migration decisions

	(1)	(2)	(3)	(4)
$IS_{i,t}$	-0.84	-0.84	-0.99	-0.98
	(1.02)	(0.98)	(1.00)	(1.03)
$Contribution_{i,t} - Contribution_{g(-i),t}$	1.45**	1.40**	1.36**	1.39**
	(0.70)	(0.67)	(0.66)	(0.69)
$Average\ Payoff_{h\neq g,t} - Average\ Payoff_{g,t}$	0.46**	0.48**	0.45**	0.49**
	(0.19)	(0.19)	(0.19)	(0.20)
$Group\ Size_{g(i),t}$	-1.00**	-1.00**	-0.94**	-0.96**
	(0.45)	(0.46)	(0.44)	(0.44)
$Vote\ Different\ from\ Group_{i,t-1}$	2.84*	2.82*	2.76*	2.70*
	(1.46)	(1.49)	(1.46)	(1.49)
$Punishment\ Received_{i,t}$	0.21**	0.20**	0.19*	0.19*
	(0.11)	(0.10)	(0.10)	(0.10)
$Punishment\ Received_{i,t} \times IS_{i,t}$	-0.01	-0.01	-0.01	-0.01
	(0.08)	(0.08)	(0.08)	(0.08)
$Different\ Institutions_t$		-0.26		-0.33
		(0.91)		(0.93)
$Migration_{i,t-1}$			1.58	1.52
			(1.20)	(1.18)
Observations	5,000	5,000	4,800	4,800
Log likelihood	-1,000.38	-1,000.32	-998.99	-998.92

Notes. The table reports the change in the probability of migrating (in percentage points) associated with a change in the covariates; units of observation are individuals-per-period. All the models include subject, period, group, and session random effects. Standard errors are presented in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Except for $IS_{i,t}$ and its interaction with *Punishment Received* $_{i,t}$, all the predictors included in the column (1) specification have a statistically significant effect on $Pr(Migration_{i,t} = 1)$. All else equal, each token increase in the difference between i's contribution and the average contribution of the other members of her group is associated with a 1.45 percentage point rise

in the probability that she moves at the end of the period. Similarly, $Pr(Migration_{i,t} = 1)$ rises by 0.46 percentage points with each token increase in the difference between the average payoff of the members of group the other group (h) and the average earnings of the members of i's group (g). The probability that i migrates also rises by 2.84 p. points when her group did not adopt i's preferred rule-set in the most recent voting round.[27] Group size, by contrast, has a negative marginal effect on the probability that the average subject moves in a given period: each additional member of i's group reduces $Pr(Migration_{i,t} = 1)$ by one percentage point. Finally, the higher the punishment received by i, the higher the probability she leaves the group, independently of whether the punishment originates from IS or FS.

Column (2) of Table 4 adds *Different Institutions*$_t$, an indicator for periods in which the two groups have different rule-sets in place, to the explanatory variables. Surprisingly, the probability of migration does not increase when the two groups implement different sanctioning institutions. Thus, it seems that subjects care more about institutions' effectiveness in generating high earnings than about the institutions *per se*. The estimated marginal effects for the other predictors remain similar to those presented in column (1).

Columns (3) and (4) of Table 4 present the results of two specifications that incorporate the lag of the dependent variable to the models of columns (1) and (2), respectively. All the substantive results discussed above remain unchanged. We also find no evidence of persistence in individuals' migration decisions: switching groups in the previous period does not have a significant effect on the propensity to migrate in the current period.

Result 3. *On average, 20% of participants migrate every period, and group size varies by 0.4 members between any two consecutive periods. Subjects are not more likely to migrate when groups implements different sanctioning institutions, relative to periods in which the two groups play under the same rule-set. Yet, when sanctioning institutions do differ across groups, a vast majority of the "migrating subjects" moves*

[27] For a similar result, see Cobo-Reyes, Katz and Meraglia (2019).

towards groups that have informal sanctions in place.

4 Conclusions

The results of our experiment support the conclusion that the type of institutions emerging in a society depends on initial conditions as well as on whether the society is open or closed. When societies are sufficiently cooperative under an informal sanctioning mechanism, the presence of fixed groups reinforces the effectiveness of these institutions and, therefore, subjects' preference for them. However, when moving in and out of groups is possible, it becomes harder to sustain cooperation. Importantly, when voting for their preferred institution, subjects consider both (*i*) current group members' behavior and (*ii*) potential newcomers' expected behavior. Overall, subjects' expectations play a key role in leading to an early adoption of centralized formal institutions. The implication is that allowing for groups to form endogenously decreases subjects' willingness to adopt informal sanctioning institutions.

When a society is characterized by low initial cooperation levels under an informal sanctioning mechanism, the dynamics are essentially the opposite. In the presence of exogenously fixed groups, subjects can only rely on centralized formal institutions to improve cooperation. When groups form endogenously, however, self-selection (in the willingness to contribute and punish free-riders) can allow for an improvement in cooperation under informal sanctioning institutions, leading to a higher percentage of subjects voting for these institutions compared to the case of fixed groups.

Results in NMT confirm the findings in Markussen, Putterman and Tyran (2014) that informal sanctions are surprisingly popular, even when competing against a formal sanctioning institution which is theoretically predicted to generate high levels of efficiency. Findings in the MT treatment, on the other hand, show that this popularity is fragile -

it is significantly reduced when migration between groups is allowed. This rhymes well with results reported in Nicklisch, Grechenig and Thöni (2016) and Markussen, Putterman and Wang (2017), who show that the popularity of decentralized sanctions - relative to a centralized alternative - declines when the information provided to potential punishers about the contribution behavior of fellow group members is imperfect (i.e. cooperators are sometimes reported as free riders, and vice versa). The effect of opening group borders to migration is somewhat similar to providing imperfect information about other people's contributions. In both cases, it becomes more difficult to predict the future behavior of fellow group members, because information about their past behavior is imperfect. These studies all suggest that formal sanctioning institutions gain importance when networks of human interaction increase in complexity and fluidity, as would typically happen when a company expands, when a city grows, and when an economy develops an increasingly sophisticated division of labor. However, the present paper distinguishes itself from previous studies by pointing to the crucial impact of initial conditions. Groups with high initial levels of cooperation responded much more strongly to open borders than other groups.

The paper also contributes to the literature on the effects of endogenous choice of institutions. Whereas prior experimental work has compared exogenous and endogenous institutions in social dilemmas (e.g. Tyran and Feld, 2006; Dal Bo, Foster and Putterman, 2010), we compare an environment with one source of endogeneity (voting by ballot) to one with two sources of endogeneity (voting by ballot and voting with feet). We find lower levels of efficiency in the latter case than in the former, suggesting that the provision of multiple different avenues for people to affect their institutional environment may not necessarily lead to better outcomes.

References

Andreoni, J., & Gee, L. K. (2012). Gun for hire: Delegated Enforcement and Peer Punishment in Public Goods Provision. *Journal of Public Economics*, 96(11), 1036–1046.

Cobo-Reyes, R., Katz, G., & Meraglia, S. (2019). Endogenous Sanctioning Institutions and Migration Patterns: Experimental Evidence. *Journal of Economic Behavior and Organization*, 158, 575–606.

Dal Bo, P., Foster, A., & Putterman, L. (2010). Institutions and Behavior: Experimental Evidence on the Effects of Democracy. *American Economic Review*, 100 (5), 2205–29.

De Stefano, V. (2016). The Rise of the Just-in-Time Workforce: On-Demand Work, Crowdwork, and Labor Protection in the Gig-Economy. *Comparative Labor Law and Policy*, 37(1), 471–504.

Dehingkar, P., & Anderson, E. (2004). People on the move. New policy challenges for increasngly mobile populations *Natural Resources Perspectives*, 92, London: Overseas Development Institute.

Dessi, R., & Piccolo, S. (2016). Merchant guilds, taxation and social capital. *European Economic Review*, 83, 90–110.

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(5), 495–511.

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

Fehr, E., & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, 90(4), 980–994.

Fernández-Val, I., & Weidner, M. (2016). Individual and Time Effects in Nonlinear Panel Models with Large N,T. *Journal of Econometrics*, 192(1), 291-312.

Fréchette, G. (2012). Session-effects in the Laboratory. *Experimental Economics*, 15(3), 485–498.

Greene, W. (2003). *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Gürerk, O., Irlenbusch, B., & Rockenbach, B. (2006). The Competitive Advantage of Sanctioning Institutions. *Science*, 312(5770), 108–111.

Gürerk, O., Irlenbusch, B., & Rockenbach, B. (2014). On Cooperation in Open Communities. *Journal of Public Economics*, 120(1), 220–230.

Harris, & J.R., Todaro, M., (1970). Migration, Unemployment and Development. *American Economic Review*, 60(1), 126–142.

Herrmann, B., Thöni, C. & Gächter, S. (2008). Antisocial Punishment Across Societies. *Science*, 319(5868), 1362–1367.

Hirschman, A. (1970). *Exit, Voice and Loyalty. Responses to Decline in Firms, Organizations and States.* Cambridge, MA: Harvard University Press.

Honoré, B., & Kesina, M. (2017). Estimation of Some Nonlinear Panel Data Models With Both Time-Varying and Time-Invariant Explanatory Variables. *Journal of Business Economic Statistics*, 35(4), 543-558.

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(1), 38–65.

Kripfganz, S., & Schwarz, C. (2019). Estimation of Linear Dynamic Panel Data Models with Time-Invariant Regressors. Forthcoming, *Journal of Applied Econometrics*.

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

Markussen, T., Putterman, L., & Wang, L. (2017). Governing Collective Action in the Face of Observational Error. Brown Economics Working Paper Series.

Nicklisch, A., Grechenig, K., & Thöni, C. (2016). Information-Sensitive Leviathans. *Journal of Public Economics*, 144(1), 1–13.

Ogilvie, S. (2011). *Institutions and European Trade. Merchant Guilds, 1000-1800.* Cambridge, UK: Cambridge University Press.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a Sword: Self-Governance Is Possible. *American Political Science Review*, 86(2), 404–417.

Poen, E. (2009). The Tobit Model with Feedback and Random Effects: A Monte-Carlo Study. CeDEx Discussion Paper eries 2009-14, University of Nottingham.

Robbett, A. (2014). Local Institutions and the Dynamics of Community Sorting. *American Economic Journal: Microeconomics*, 6(3), 136–156.

Smith, A. (2013). Estimating the Causal Effect of Beliefs on Contributions in Repeated Public Goods. *Experimental Economics*, 16(3), 414-425.

Sutter, M., Haigner, S., & Kocher, M. G. (2010). Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. *Review of Economic Studies*, 77(4), 1540–1566.

Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, 64(5), 416–424.

Tyran, J.-R., & Feld, L.P. (2006). Achieving Compliance when Legal Sanctions are Non-deterrent. *Scandinavian Journal of Economics*, 108(1), 135–156.

Weber, R. (2006). Managing Growth to Achieve Efficient Coordination in Large Groups. *American Economic Review*, 96(1), 114—126.

Wooldridge, J. (2005). Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. *Journal of Applied Econometrics*, 20(1), 39–54.

Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Zhang, B., Li, C., De Silva, H., Bednarik, P., & Sigmund, K. (2014). The Evolution of Sanctioning Institutions: an Experimental Approach to the Social Contract. *Experimental Economics*, 17(2), 285–303.

Appendix

A. Experimental Instructions

Moving Treatment (MT)

You are now taking part in an economic experiment. Depending on your decisions and the decisions of other participants, you will be able to earn money. These instructions describe how you can earn money. Please read them carefully.

During the experiment you are not allowed to communicate with other participants. If you have a question, please raise your hand. One of us will come to answer your question. Sometimes you may have to wait a short while before the experiment continues. Please be patient. There will be a total of 30 periods in this experiment. We will explain carefully what you have to do.

Allocations: In each period, you will be in a group of some size (you could also be by yourself). In each period, all members of your group, including you, will receive 50 points as "endowment".

You and the four others in your group simultaneously decide how to use the endowment. There are two possibilities:

1. Allocate points to a private account
2. Allocate points to a group account.

You will be asked to indicate the number of points you want to allocate to the group account. Only integers between 0 and 50 are allowed for this purpose. The remaining points will automatically be allocated to your private account.

Your payoff from the allocation decisions of yourself and the other members of your group will be as follows:

50 – (points you allocate to the group account) + $(1.6/n)\times$ (sum of points allocated by all members in the group to the group account)

where n is the number of members in your group. If you are the only member in a group, your payoff in that period will be 50 points.

EXAMPLE: Consider the case in which one participant belongs to a group in a particular period. Assume that the group has four members. For each point the participant allocates to her private account, she earns one point. For each point she allocates to the group account, she earns 0.4 points and each other group member also earns 0.4 points. If the total number of points in the group account is 100, the participant receives a payoff of (100 1.6)/4=40 points from the group account (plus the number of points she has allocated to her\his private account).

Groups: In the first period you will be assigned to a group of five people (including yourself). You can be assigned to either GROUP A or GROUP B. Groups will remain fixed for the first 5 periods. Starting in period 6, you will decide at the end of each period whether you want to move to the other group. The other members of your group and the other group will also choose whether to move. We will now explain in more detail how this works.

Payment rules: there will be **two** different **Rule Sets**, which affect your earnings in different ways:

RULE SET 1 (peer to peer points reduction): In Rule Set 1, there are two stages in each period. In the first stage, you make your allocation decision as described above and learn the decisions of the other group members along with your earnings. In the second stage, you have an opportunity to reduce the earnings of others in your group at a cost to you. Here is how it works.

After the first stage of each period, you will be shown the amount allocated to the group account by each of the others in your group, in a random order, and in a box below that information you will be asked to enter a number of points (if any) that you wish to use to reduce the earnings of the individual who made that allocation decision (see below). Each point you allocate to reducing another's earnings reduces your own earnings by 1 point and reduces that individual's earnings by 3 points. Your own earnings can be reduced in the same way by the decisions of others in your group.

You are free to leave any or all of the others' earnings unchanged by entering 0s in the relevant boxes. Your payoffs under Rule set 1 will be computed as follows:

50 – (points you allocate to group account) + $(1.6/n)\times$
(sum of points allocated by all in the group to the group account) −
(points you spend to reduce others' earnings) −
3 × (sum of reduction points directed at you by others in your group)

EXAMPLE: Consider the case in which one participant belongs to a group of 5 people and RULE SET 1 (peer to peer points reduction) is in place. Suppose that you use 0 points to reduce the earnings of the first and second group members, you use 1 point to reduce the earnings of the third, and you use 2 points to reduce the earnings of the fourth. Suppose further that these individuals use 0, 1, 0 and 3 points, respectively, to

reduce your earnings. Then, the third and fourth individuals' earnings for the period will be reduced by 3 and by 6 points, respectively, in addition to any reductions due to the decisions of others. Your own earnings for the period will be reduced by 3 points = the cost for you of imposing reductions on others, plus $(1 \times 3) + (3 \times 3) = 12$ points = the reductions imposed on your earnings by others. At the end of the reduction stage, you will learn that your earnings were reduced by others by a total of 12 points, but you will not be told which individuals reduced your earnings or by how much any given individual reduced your earnings. Others will also not know who reduced their earnings.

The earnings reduction process is subject to two limits. First, you cannot assign more than 10 reduction points to any one individual in your group. Second, the total effective reduction of your earnings due to others' decisions in a given period cannot be greater than your total earnings from the allocation stage of that period. For example, if your earnings after the allocation stage are 26 points and if others use a total of 9 points to reduce your earnings, you will lose only 26 points, not $9 \times 3 = 27$. However, the points that you spend to reduce the earnings of others are always costly to you, even if that brings your earnings for a period to less than zero. To continue with the example in which you earn 26 points before reductions and you lose 26 points due to the (27 points worth of) reductions others impose on you: if in the same period you have chosen to spend 3 points on reducing others' earnings, your total earnings for the period are – 3. Points lost in some periods are deducted from the £3 show-up fee that you will receive at the end of the experiment.

RULE SET 2 (automatic points reduction): In Rule Set 2, each individual pays a fixed fee of 5 points in each period. The fee is deducted from your earnings at the end of the period. In addition, each individual pays a fine equal to 80 percent of the amount of points she\he allocates to the private account in that period. Payoffs in each period are calculated as

follows:

50 – (points you allocate to the group account) $+ (1.6/n)\times$
(sum of points allocated by all members in the group to the group account) $-$
$0.8 \times$ (points you allocate to the private account) $- 5$

Note that if you are the only member in a group and Rule Set 2 applies, your payoff in that period will be 45 points (50 points minus the 5 points that you pay as fixed fee).

EXAMPLE: Consider the case in which one participant belongs to a group of 5 people and RULE SET 2 (automatic points reduction) is in place. If the participant contributes 20 points to the group account and keeps 30 points for the private account, given RULE SET 2, that participant would have to pay a fine of $0.8 \times 30 = 24$ points.

Payment rules and groups in the first 5 periods: If you are assigned to Group A, you will remain in Group A for the first 5 periods. If you are assigned to Group B you will stay in Group B for the first 5 periods. Rule set 1 (peer to peer points reductions) will apply in both groups for the first 5 periods.

Information about earnings and rule sets: At the end of each period, you will observe: (*i*) the average final earnings of your current group, (*ii*) the average final earnings of the other group, (*iii*) the rule set of your current group, and (*iv*) the rule set of the other group.

Moving between groups: Starting in period 6, at the end of each period you will decide whether you wish to leave the group you are currently in and move to the other group. The other members of your group, and the members of the other group, will make the same decision. You and all other participants are allowed to move back and forth between groups as many times as you wish. Hence, the size and member composition of the groups may change from period to period.

Voting for rule sets: Initial rules for Groups A and B will apply only for the first five periods. From period 6, every five periods you will have the chance to vote on whether you want Rule Set 1 or Rule Set 2 to be implemented in the group you currently belong to. Note that when you vote, you will vote to establish a rule in the group to which you belong in the voting period. Whichever rule set receives the highest number of votes will be in effect for the following five periods.

Payment: At the end of the experiment points will be converted into GBP at the rate of 15 points = 1GBP. You will be paid only for 3 periods, randomly chosen.

SUMMARY

1. There will be a total of 30 periods
2. You will begin by being in a group of five persons, including you.
3. In each period, you will allocate 50 points between a private and a group account.
4. Starting in period 6, you and everybody else can decide at the end of each period whether to move to the other group or not.
5. There are two Rule Sets. In Rule Set 1, each group member can reduce other group members' earnings after seeing the allocations of each individual to the group account.

It costs 1 point to reduce the earnings of another group member by 3 points. In Rule Set 2, there is a fixed cost of 5 points in each period, deducted from your earnings at the end of the period. Each individual pays a fine equal to 80 percent of the amount of points he or she allocated to the private account. Every five periods you will vote on which of the two rule sets you prefer.

6. You will be paid only for 3 periods (randomly chosen).

No Moving Treatment (NMT)

You are now taking part in an economic experiment. Depending on your decisions and the decisions of other participants, you will be able to earn money. These instructions describe how you can earn money. Please read them carefully.

During the experiment you are not allowed to communicate with other participants. If you have a question, please raise your hand. One of us will come to answer your question. Sometimes you may have to wait a short while before the experiment continues. Please be patient.

There will be a total of 30 periods in this experiment. We will explain carefully what you have to do.

Allocations: In each period, you will be in a group of 5 people. In each period, all members of your group, including you, will receive 50 points as "endowment".

You and the four others in your group simultaneously decide how to use the endowment. There are two possibilities:

1. Allocate points to a private account
2. Allocate points to a group account.

You will be asked to indicate the number of points you want to allocate to the group account. Only integers between 0 and 50 are allowed for this purpose. The remaining points will automatically be allocated to your private account.

Your payoff from the allocation decisions of yourself and the other members of your group will be as follows:

50 – (points you allocate to the group account) + (1.6/5)×
(sum of points allocated by all members in the group to the group account)

EXAMPLE: Consider the case in which one participant belongs to a group in a particular period. Assume that the group has five members. For each point the participant allocates to her private account, she earns one point. For each point she allocates to the group account, she earns 0.32 points and each other group member also earns 0.32 points. If the total number of points in the group account is 100, the participant receives a payoff of $(100 \times 1.6)/5 = 32$ points from the group account (plus the number of points she has allocated to her\ his private account).

Groups: In the first period you will be assigned to a group of five people (including yourself). You can be assigned to either GROUP A or GROUP B. Your group will remain the same until the end of the experiment.

Payment rules: there will be **two** different **Rule Sets**, which affect your earnings in different ways:

RULE SET 1(peer to peer points reduction): In Rule Set 1, there are two stages in each period. In the first stage, you make your allocation decision as described above and learn

the decisions of the other group members along with your earnings. In the second stage, you have an opportunity to reduce the earnings of others in your group at a cost to you. Here is how it works.

After the first stage of each period, you will be shown the amount allocated to the group account by each of the others in your group, in a random order, and in a box below that information you will be asked to enter a number of points (if any) that you wish to use to reduce the earnings of the individual who made that allocation decision (see below). Each point you allocate to reducing another's earnings reduces your own earnings by 1 point and reduces that individual's earnings by 2 points. Your own earnings can be reduced in the same way by the decisions of others in your group.

You are free to leave any or all of the others' earnings unchanged by entering 0s in the relevant boxes. Your payoffs under Rule set 1 will be computed as follows:

50 – (points you allocate to group account) +
(1.6/5)×
(sum of points allocated by all in the group to the group account) −
(points you spend to reduce others' earnings) −
2 × (sum of reduction points directed at you by others in your group)

EXAMPLE: Consider the case in which one participant belongs to a group of 5 people and RULE SET 1 (peer to peer points reduction) is in place. Suppose that you use 0 points to reduce the earnings of the first and second group members, you use 1 point to reduce the earnings of the third, and you use 2 points to reduce the earnings of the fourth. Suppose further that these individuals use 0, 1, 0 and 3 points, respectively, to reduce your earnings. Then, the third and fourth individuals' earnings for the period will

be reduced by 2 and by 4 points, respectively, in addition to any reductions due to the decisions of others. Your own earnings for the period will be reduced by 3 points = the cost for you of imposing reductions on others, plus $(1 \times 2) + (3 \times 2) = 8$ points = the reductions imposed on your earnings by others. At the end of the reduction stage, you will learn that your earnings were reduced by others by a total of 8 points, but you will not be told which individuals reduced your earnings or by how much any given individual reduced your earnings. Others will also not know who reduced their earnings.

The earnings reduction process is subject to two limits. First, you cannot assign more than 10 reduction points to any one individual in your group. Second, the total effective reduction of your earnings due to others' decisions in a given period cannot be greater than your total earnings from the allocation stage of that period. For example, if your earnings after the allocation stage are 26 points and if others use a total of 14 points to reduce your earnings, you will lose only 26 points, not $14 \times 2 = 28$. However, the points that you spend to reduce the earnings of others are always costly to you, even if that brings your earnings for a period to less than zero. To continue with the example in which you earn 26 points before reductions and you lose 26 points due to the (28 points worth of) reductions others impose on you: if in the same period you have chosen to spend 3 points on reducing others' earnings, your total earnings for the period are – 3. Points lost in some periods are deducted from the £3 show-up fee that you will receive at the end of the experiment.

RULE SET 2 (automatic points reduction): In Rule Set 2, each individual pays a fixed fee of 5 points in each period. The fee is deducted from your earnings at the end of the period. In addition, each individual pays a fine equal to 80 percent of the amount of points sheallocates to the private account in that period. Payoffs in each period are calculated as follows:

50 – (points you allocate to the group account) + (1.6/5)×
(sum of points allocated by all members in the group to the group account) −
0.8 × (points you allocate to the private account) − 5

Note that if you are the only member in a group and Rule Set 2 applies, your payoff in that period will be 45 points (50 points minus the 5 points that you pay as fixed fee).

EXAMPLE: Consider the case in which one participant belongs to a group of 5 people and RULE SET 2 (automatic points reduction) is in place. If the participant contributes 20 points to the group account and keeps 30 points for the private account, given RULE SET 2, that participant would have to pay a fine of $0.8 \times 30 = 24$ points.

Payment rules in the first 5 periods: Rule set 1 (peer to peer points reductions) will apply in both groups for the first 5 periods.

Information about earnings and rule sets: At the end of each period, you will observe: (*i*) the average final earnings of your current group, (*ii*) the average final earnings of the other group, (*iii*) the rule set of your current group, and (*iv*) the rule set of the other group.

Voting for rule sets: Initial rules for Groups A and B will apply only for the first five periods. From period 6, every five periods you will have the chance to vote on whether you want Rule Set 1 or Rule Set 2 to be implemented in your group. Whichever rule set receives the highest number of votes will be in effect for the following five periods.

Payment: At the end of the experiment tokens will be converted into GBP at the rate of 15 points = 1GBP. You will be paid only for 3 periods randomly chosen.

SUMMARY

1. There will be a total of 30 periods
2. You will begin by being in a group of five persons, including you.
3. In each period, you will allocate 50 points between a private and a group account.
4. There are two Rule Sets. In Rule Set 1, each group member can reduce other group members' earnings after seeing the allocations of each individual to the group account. It costs 1 point to reduce the earnings of another group member by 2 points. In Rule Set 2, there is a fixed cost of 5 points in each period, deducted from your earnings at the end of the period. Each individual pays a fine equal to 80 percent of the amount of points he or she allocated to the private account. Every five periods you will vote on which of the two rule sets you prefer.
5. You will be paid only for 3 periods (randomly chosen).

B. Additional Results

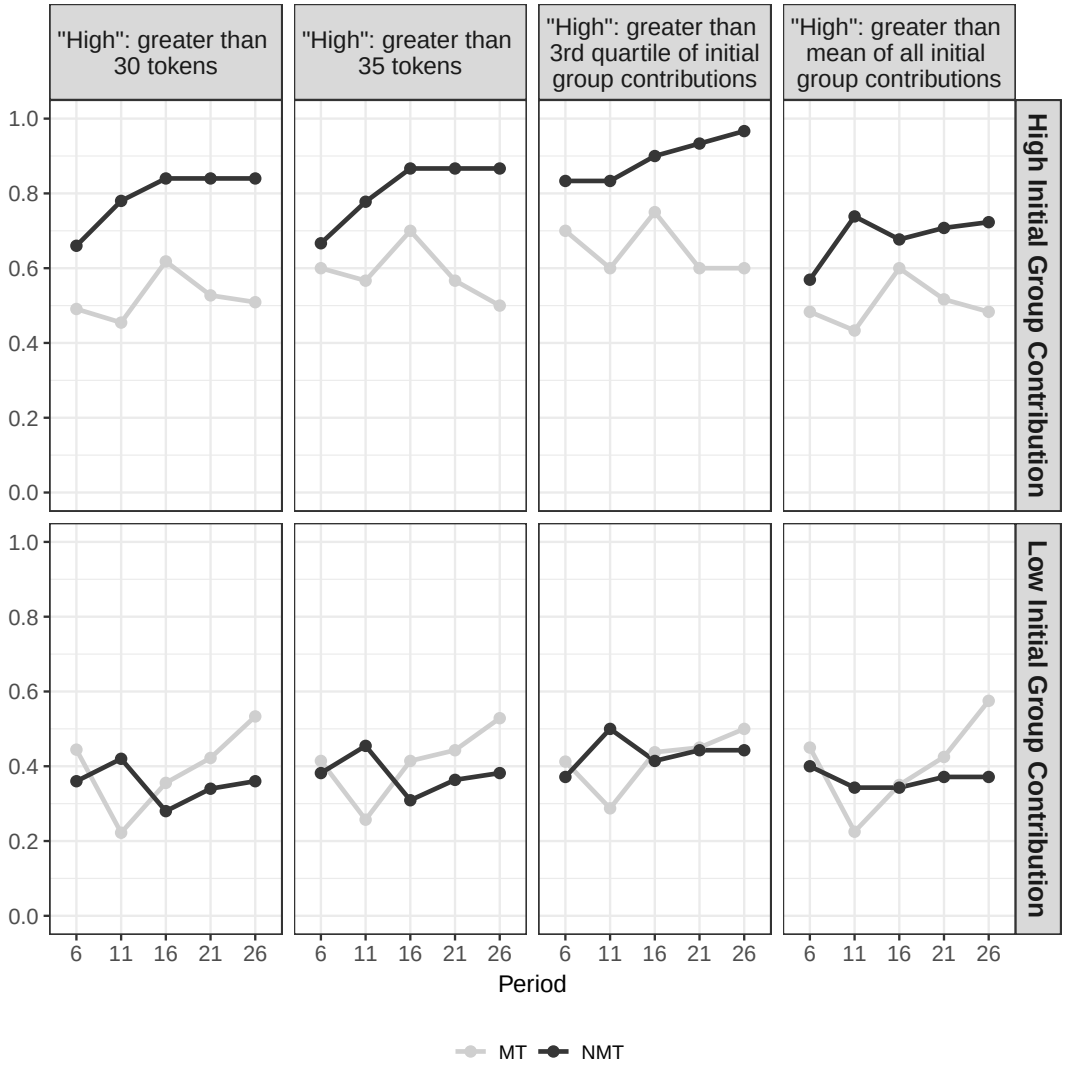

Figure B.1: Impact of initial group contributions on voting patterns under alternative definitions of "high" initial group contributions. The figure plots the percentage of subjects voting for IS over time under MT and NMT, distinguishing groups by their average contribution in the first five periods of the experiment. "High" contributions are operationalized by four alternative measures. In the first two (leftmost) vertical panels, groups with (relatively) "high" initial contributions are taken as those whose members' average contribution in the first five periods of the experiment are greater than 30/35 tokens. In the third panel (column), groups with (relatively) "high" initial contributions are defined as those whose members' average initial contribution exceeds the third quartile of the distribution of initial contributions across all the groups in the experiment. Finally, in the rightmost panel, groups with (relatively) "high" initial contributions are those whose members' average contribution in the first five periods is greater than the mean of all groups' initial contributions.

Table B.1
Parameter estimates
Panel probit models for voting

	(1)	(2)	(3)
Intercept	-0.54	-0.99***	-0.89***
	(0.35)	(0.32)	(0.30)
$Moving_i$	0.49	0.49	0.41
	(0.39)	(0.35)	(0.32)
High Initial Group Contribution$_{g(i)}$	1.39***	1.12***	1.00***
	(0.30)	(0.30)	(0.28)
High Initial Group Contribution$_{g(i)} \times Moving_i$	-1.10***	-0.94**	-0.82**
	(0.38)	(0.35)	(0.34)
High Initial Group Payoff$_{h \neq g}$	0.16	0.23	0.19
	(0.27)	(0.24)	(0.23)
High Initial Group Payoff$_{h \neq g} \times Moving_i$	-0.41	-0.36	-0.29
	(0.33)	(0.33)	(0.32)
$Contribution_{i,t-1} - Contribution_{g(-i),t-1}$	-0.25**	-0.17	-0.18
	(0.10)	(0.12)	(0.12)
$IS_{i,t-1}$	0.57***	0.35**	0.43***
	(0.14)	(0.15)	(0.15)
Group Size$_{g(i),t}$	-0.05	-0.03	-0.04
	(0.03)	(0.03)	(0.03)
$Vote_{i,t-1}$		0.91***	0.99***
		(0.14)	(0.14)
$Payoff_{i,t-1}$			-0.66**
			(0.26)
$Payoff_{i,t-1} \times IS_{i,t-1}$			1.81***
			(0.37)
Observations	1,000	800	800
Log likelihood	-549.47	-393.09	-378.89

Notes. The table reports "raw" parameter estimates for the panel probit models used to compute the marginal effects presented in Table 1. Units of observation are individuals-per-period, with the sample restricted to the five periods that include a voting stage. All specifications include subject, period, group, and session random effects. Standard errors are presented in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table B.2
Additional specifications examining individual voting decisions

	(1)	(2)	(3)
$Moving_i$	11.59	13.17	10.90
	(10.12)	(9.43)	(8.09)
$High\ Initial\ Group\ Payoff_{g(i)}$	33.56***	25.79***	23.70***
	(6.66)	(6.85)	(6.28)
$High\ Initial\ Group\ Payoff_{g(i)} \times Moving_i$	-19.36*	-19.76*	-16.40*
	(11.33)	(10.28)	(9.03)
$High\ Initial\ Group\ Payoff_{h \neq g}$	8.77	10.49	9.17
	(6.83)	(6.50)	(5.78)
$High\ Initial\ Group\ Payoff_{h \neq g} \times Moving_i$	-13.39	-11.15	-9.18
	(10.13)	(9.74)	(8.80)
$Contribution_{i,t-1} - Contribution_{g(-i),t-1}$	-0.56**	-0.36	-0.18
	(0.24)	(0.24)	(0.12)
$IS_{i,t-1}$	18.19***	10.91**	12.20***
	(4.17)	(4.26)	(4.39)
$Group\ Size_{g(i),t}$	-1.47	-0.93	-1.20
	(0.97)	(0.89)	(0.88)
$Vote_{i,t-1}$		28.42***	28.70***
		(4.85)	(4.74)
$Payoff_{i,t-1}$			-0.52***
			(0.19)
$Payoff_{i,t-1} \times \quad IS_{i,t-1}$			2.54***
			(0.47)
Observations	1,000	800	800
Log likelihood	-550.96	-395.27	-379.02

Notes. The table replicates the analyses in Table 1 of the paper, replacing *High Initial Group Contribution*$_{g(i)}$ with *High Initial Group Payoff*$_{g(i)}$ as a predictor of $Pr(Vote_{i,t} = 1)$. Units of observation are individuals-per-period, with the sample restricted to the five periods that include a voting stage. All specifications include subject, period, group, and session random effects. Standard errors are presented in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table B.3
Parameter estimates
Fixed effects models for voting

	(1)	(2)	(3)
Intercept	0.26***	0.18***	0.21***
	(0.06)	(0.06)	(0.06)
$Moving_i$	0.04	0.03	0.03
	(0.04)	(0.04)	(0.04)
$High\ Initial\ Group\ Payoff_{g(i)}$	0.28***	0.14***	0.11**
	(0.04)	(0.04)	(0.04)
$High\ Initial\ Group\ Payoff_{g(i)} \times Moving_i$	-0.20***	-0.11*	-0.10*
	(0.06)	(0.06)	(0.06)
$High\ Initial\ Group\ Payoff_{h \neq g}$	0.04	0.07	0.06
	(0.05)	(0.05)	(0.05)
$High\ Initial\ Group\ Payoff_{h \neq g} \times Moving_i$	-0.03	-0.04	-0.03
	(0.05)	(0.05)	(0.05)
$Contribution_{i,t-1} - Contribution_{g(-i),t-1}$	-0.37***	-0.37***	-0.35***
	(0.12)	(0.12)	(0.12)
$IS_{i,t-1}$	0.33***	0.33***	0.32***
	(0.03)	(0.03)	(0.03)
$Group\ Size_{g(i),t}$	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)
$Vote_{i,t-1}$		0.31***	0.31***
		(0.03)	(0.03)
$Payoff_{i,t-1}$			-0.28***
			(0.07)
$Payoff_{i,t-1} \times \ IS_{i,t-1}$			0.60***
			(0.09)
Observations	1,000	800	800

Notes. The table reports parameter estimates for (linear) fixed-effects panel models for voting obtained using the two-step approach proposed by Kripfganz and Schwarz (2019) to identify the coefficients of time-invariant regressors. Units of observation are individuals-per-period, with the sample restricted to the five periods that include a voting stage. Standard errors are presented in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

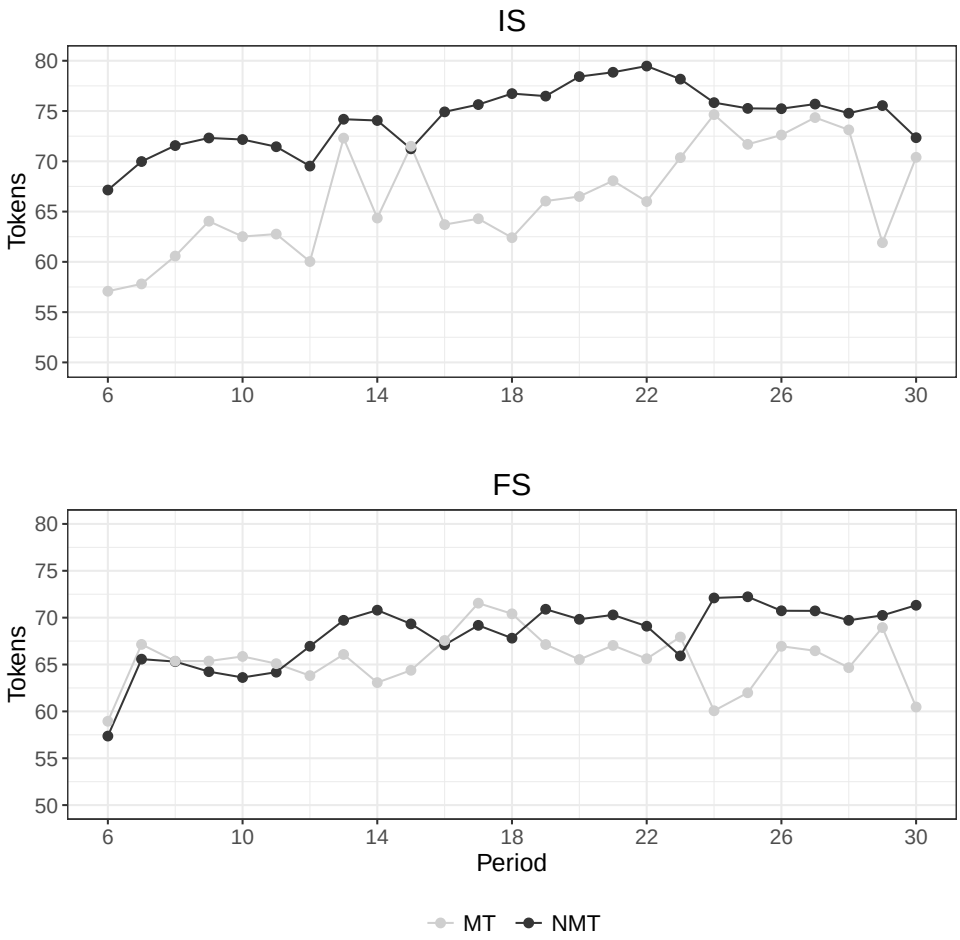

Figure B.2: Average individual payoffs per period over time, by rule-set and treatment. The figure complements the information displayed in Figure 4, plotting the average individual payoff per period in groups implementing informal and formal sanctioning institutions under both treatments. Gray (black) lines represent the average individual payoff per period under MT (NMT) for subjects playing under IS (upper panel) and FS (lower panel), averaged across sessions. As in the case of contributions, the average payoff under IS is significantly higher in NMT (74.41 tokens) than in MT (66.86 tokens) (Mann-Whitney test: $z = 2.04$, $p = 0.04$, two-tailed). By contrast, differences between treatments are not statistically significant under FS (Mann-Whitney test: $z = 0.98$, $p = 0.33$, two-tailed). See also Table B.4 below.

Table B.4
Difference in average individual payoffs between rule-sets, by treatment

Treatment	Rule-set	Average individual payoff (per-period, in tokens)	Wilcoxon signed rank test (two-sided)
MT	IS	60.63	$z = -2.29$, $p = 0.02$
	FS	65.59	
NMT	IS	67.59	$z = -0.41$, $p = 0.74$
	FS	68.29	

Notes. Like in the case of contributions, we observe that differences in payoffs are driven by the behavior of subjects in open societies: while in MT the average individual payoff per-period is significantly higher under FS than under IS, the difference between the two institutional settings is statistically indistinguishable from zero in NMT.

Table B.5
Parameter estimates
Censored regression models for contributions

	(1)	(2)
Intercept	7.77***	3.00
	(2.50)	(2.56)
$Moving_i$	-1.28	-3.03
	(2.82)	(2.67)
$IS_{i,t}$	-13.66***	-13.32***
	(1.79)	(1.78)
$Moving_i \times IS_{i,t}$	-5.08**	-4.90**
	(2.04)	(2.00)
High Initial Group Contribution$_{g(i)}$	5.18**	4.04*
	(2.18)	(2.15)
$Contribution_{g(-i),t-1}$	0.85***	0.97***
	(0.09)	(0.09)
Group Size$_{g(i),t}$	1.38***	0.64*
	(0.35)	(0.35)
$Contribution_{i,t-1}$	0.39***	0.46***
	(0.08)	(0.08)
Punishment Received$_{i,t-1}$	-0.12	-0.01
	(0.10)	(0.10)
$IS_{i,t-1}$	7.25***	8.88***
	(1.57)	(1.65)
Punishment Received$_{i,t-1} \times IS_{i,t-1}$	0.22**	
	(0.10)	
Punishment Received$_{i,t-1} \times \mathrm{IS}_{i,t-1} \times$		0.49
$\mathbb{1}\big(Contribution_{i,t-1} > Contribution_{g(-i),t-1}\big)$		(0.33)
Punishment Received$_{i,t-1} \times \mathrm{IS}_{i,t-1} \times$		0.16*
$\mathbb{1}\big(Contribution_{i,t-1} < Contribution_{g(-i),t-1}\big)$		(0.09)
Observations	4,800	4,800
Log likelihood	-11,558.33	-11,368.93

Notes. The table reports "raw" parameter estimates for the panel data (doubly) censored regression models used to compute the marginal effects presented in Table 2. Units of observation are individuals-per-period. All the specifications include subject, period, group, and session random effects. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table B.6
Parameter estimates
Fixed-effects censored regression models for contributions

	(1)	(2)
$Moving_i$	-1.74	-2.88
	(3.01)	(3.02)
$IS_{i,t}$	-26.07***	-25.19***
	(2.56)	(2.51)
$Moving_i \times IS_{i,t}$	-6.73**	-5.35**
	(2.74)	(2.70)
$High\ Initial\ Group\ Contribution_{g(i)}$	6.21**	5.02**
	(2.56)	(2.55)
$Contribution_{g(-i),t-1}$	0.93***	1.08***
	(0.09)	(0.09)
$Group\ Size_{g(i),t}$	2.19***	1.70***
	(0.30)	(0.30)
$Contribution_{i,t-1}$	0.20***	0.33***
	(0.08)	(0.08)
$Punishment\ Received_{i,t-1}$	-0.15	-0.05
	(0.11)	(0.10)
$IS_{i,t-1}$	11.56***	15.75***
	(1.98)	(2.04)
$Punishment\ Received_{i,t-1} \times IS_{i,t-1}$	0.18*	
	(0.10)	
$Punishment\ Received_{i,t-1} \times \mathrm{IS}_{i,t-1} \times$		0.46
$\mathbb{1}\big(Contribution_{i,t-1} > Contribution_{g(-i),t-1}\big)$		(0.33)
$Punishment\ Received_{i,t-1} \times \mathrm{IS}_{i,t-1} \times$		0.17*
$\mathbb{1}\big(Contribution_{i,t-1} < Contribution_{g(-i),t-1}\big)$		(0.10)
Observations	4,800	4,800

Notes. The table reports parameter estimates for fixed-effects censored regression models for contributions using the two-step approach proposed Honoré and Kesina (2017) to identify the coefficients of time-invariant predictors. Units of observation are individuals-per-period. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table B.7
Parameter estimates
Censored regression models for punishment

	(1)	(2)
Intercept	14.98***	6.93***
	(1.79)	(2.37)
$Moving_i$	6.24**	1.73
	(2.53)	(2.60)
$Contribution_{i,t}$	-1.01***	-0.82***
	(0.04)	(0.07)
$Contribution_{i,t} \times Moving_i$		0.18**
		(0.08)
$max(Contribution_{g(-i),t}$ - $Contribution_{i,t}, 0)$		0.47***
		(0.07)
$max(Contribution_{i,t}$ - $Contribution_{g(-i),t}, 0)$		0.04
		(0.07)
Observations	2,399	2,399
Log likelihood	-4,573.076	-4,394.99

Notes. The table reports "raw" parameter estimates for the panel data (doubly) censored regression models used to compute the marginal effects presented in Table 3. Units of observation are individuals-per-period, with the sample restricted to subjects playing under IS. All the specifications include subject, period, group, and session random effects. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table B.8
Parameter estimates
Fixed-effects censored regression models for punishment

	(1)	(2)
$Moving_i$	7.07**	3.54
	(3.08)	(3.03)
$Contribution_{i,t}$	-1.12***	-0.98***
	(0.05)	(0.09)
$Contribution_{i,t} \times Moving_i$		0.23**
		(0.11)
$max(Contribution_{g(-i),t}$ - $Contribution_{i,t}, 0)$		0.68***
		(0.09)
$max(Contribution_{i,t}$ - $Contribution_{g(-i),t}, 0)$		0.09
		(0.09)
Observations	2,399	2,399

Notes. The table reports parameter estimates for fixed-effects censored regression models for punishment received using the two-step approach proposed Honoré and Kesina (2017) to identify the coefficients of time-invariant predictors. Units of observation are individuals-per-period, with the sample restricted to subjects playing under IS. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

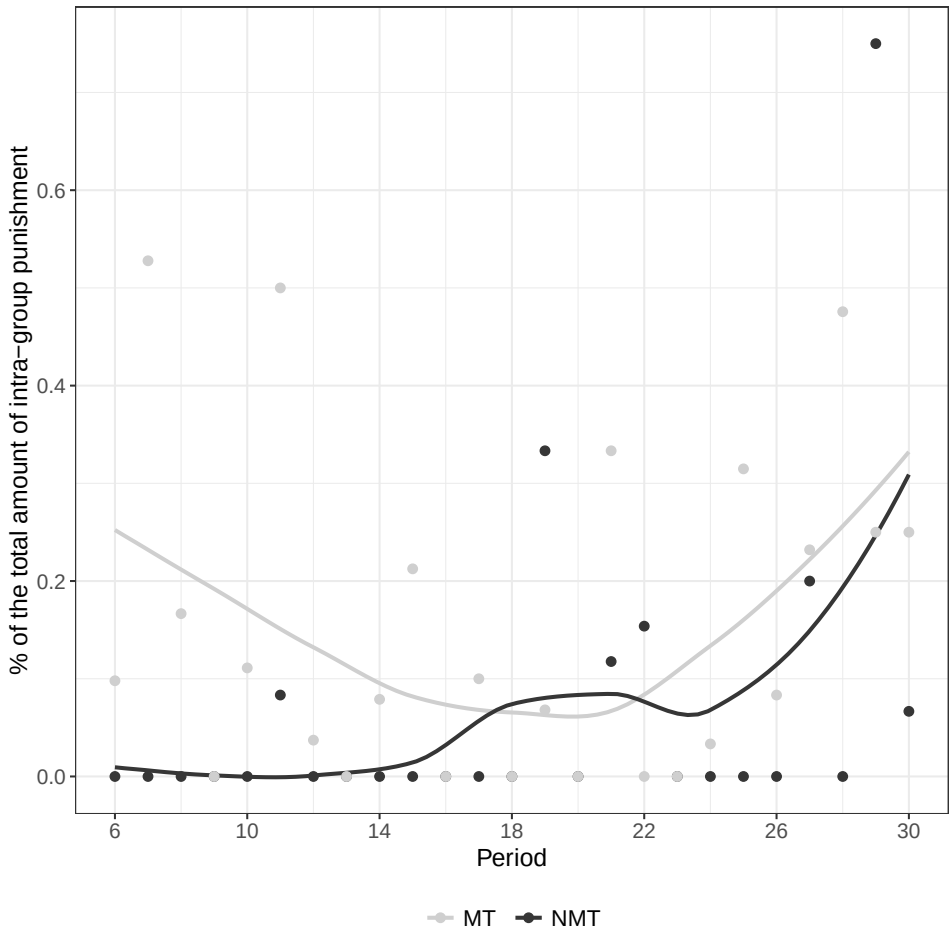

Figure B.3: Share of perverse punishment in IS, by period and treatment. The circles represent the share of "perverse" or anti-social punishment - defined as the proportion of the punishment directed towards subjects contributing more than the group average - per period in groups implementing informal sanctioning institutions under the two treatments. Solid lines represent locally weighted scatterplot smoothing curves fitted to the data.

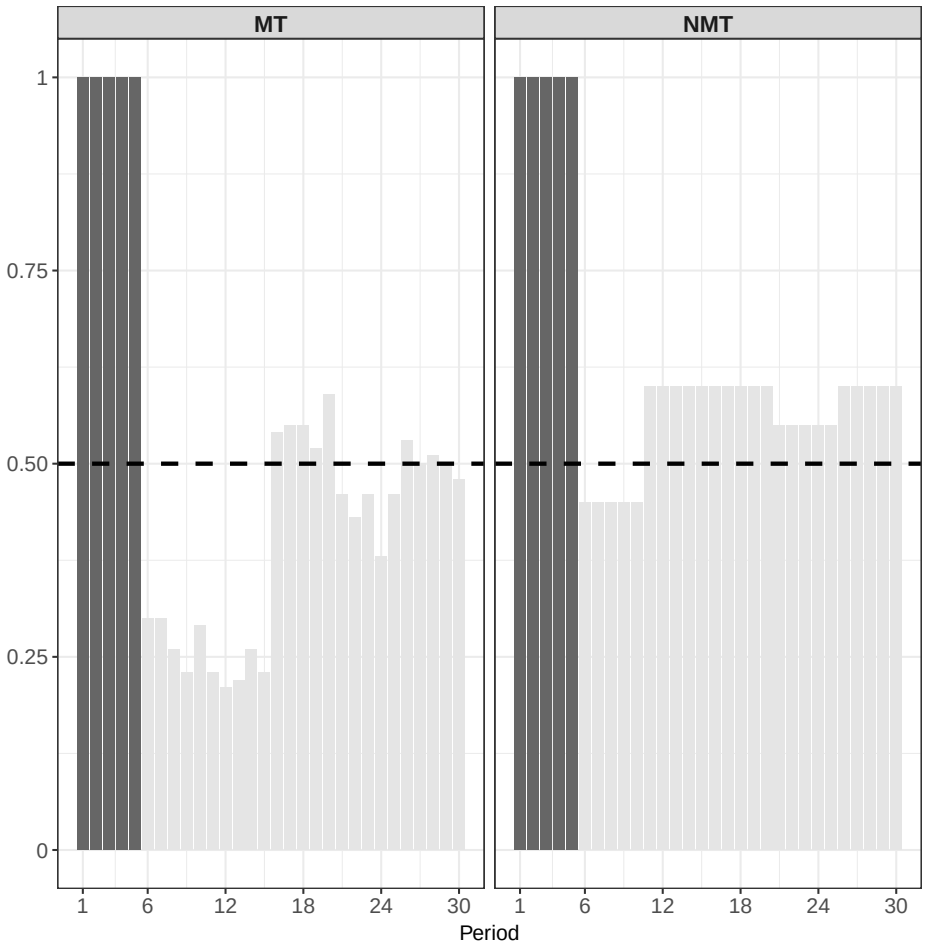

Figure B.4: Share of participants in groups implementing IS, by period and treatment. The bars represent the proportion of subjects located in groups with informal sanctioning institutions in each period under MT (left panel) and NMT (right panel), averaged across sessions. Darker bars correspond to the first 5 periods of the experiment, when groups are fixed and they all implement informal sanctions; lighter bars start in period 6, once subjects are allowed to move between groups and to vote – every 5 periods – on the rule-set to be implemented in their current group (see Section 2). The dashed horizontal line corresponds to the situation in which participants are equally distributed between IS and FS.

Table B.9
Parameter estimates
Panel probit models for migration

	(1)	(2)	(3)	(4)
Intercept	-4.63***	-4.57***	-4.61***	-4.55***
	(0.72)	(0.72)	(0.73)	(0.73)
$IS_{i,t}$	-0.09	-0.08	-0.10	-0.10
	(0.08)	(0.08)	(0.08)	(0.08)
$Contribution_{g(i),t-1} - Contribution_{g(-i),t-1}$	1.69***	1.68***	1.63***	1.63***
	(0.39)	(0.39)	(0.39)	(0.39)
$Average\ Payoff_{h\neq g,t} - Average\ Payoff_{g,t}$	0.94***	0.98***	0.94***	0.98***
	(0.08)	(0.09)	(0.08)	(0.09)
$Group\ Size_{g(i),t}$	-0.10***	-0.10***	-0.10***	-0.10***
	(0.02)	(0.02)	(0.02)	(0.02)
$Vote\ Different\ from\ Group_{i,t-1}$	0.27***	0.26***	0.26***	0.26***
	(0.09)	(0.09)	(0.09)	(0.09)
$Punishment\ Received_{i,t}$	0.38***	0.38***	0.36***	0.36***
	(0.11)	(0.11)	(0.11)	(0.11)
$Punishment\ Received_{i,t} \times IS_{i,t}$	-0.01	-0.01	-0.01	-0.01
	(0.12)	(0.12)	(0.12)	(0.12)
$Different\ Institutions_t$		-0.03		-0.03
		(0.08)		(0.08)
$Migration_{i,t-1}$			0.14	0.14
			(0.09)	(0.09)
Observations	5,000	5,000	4,800	4,800
Log likelihood	-1,000.38	-1,000.32	-998.99	-998.92

Notes. The table reports "raw" parameter estimates for the panel probit models used to compute the marginal effects presented in Table 4. Units of observation are individuals-per-period. All the models include subject, period, group, and session random effects. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table B.10
Parameter estimates
Fixed-effects panel probit models for migration

	(1)	(2)	(3)	(4)
$IS_{i,t}$	0.12	0.08	0.10	0.07
	(0.09)	(0.09)	(0.09)	(0.09)
$Contribution_{g(i),t-1} - Contribution_{g(-i),t-1}$	1.73***	1.74***	1.70***	1.70***
	(0.44)	(0.44)	(0.44)	(0.44)
$Average\ Payoff_{h\neq g,t} - Average\ Payoff_{g,t}$	0.70***	0.70***	0.71***	0.70***
	(0.09)	(0.09)	(0.09)	(0.09)
$Group\ Size_{g(i),t}$	-0.14***	-0.15***	-0.14***	-0.14***
	(0.02)	(0.02)	(0.02)	(0.02)
$Vote\ Different\ from\ Group_{i,t-1}$	0.28***	0.28***	0.28***	0.27***
	(0.09)	(0.09)	(0.09)	(0.09)
$Punishment\ Received_{i,t}$	0.31**	0.31**	0.29**	0.29**
	(0.12)	(0.12)	(0.12)	(0.12)
$Punishment\ Received_{i,t} \times IS_{i,t}$	0.10	0.08	0.10	0.08
	(0.12)	(0.12)	(0.12)	(0.12)
$Different\ Institutions_t$		0.02		0.03
		(0.09)		(0.09)
$Migration_{i,t-1}$			0.11	0.12
			(0.09)	(0.09)
Observations	2,325	2,325	2,325	2,325
Log likelihood	-854.05	-853.65	-853.84	-853.45

Notes. The table reports parameter estimates from bias-corrected fixed-effects panel probit models for migration using Fernández-Val and Weidner (2016)'s method. The bias correction is obtained from jackknife estimates; using an analytical correction yields virtually identical results. Differences in the number of observations vis-à-vis Tables 4 and B.9 are due to the fact that Fernández-Val and Weidner (2016)'s method drops subjects for whom the dependent variable does not change over time. All the models include subject and period fixed-effects. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.