Endogenous Sanctioning Institutions and Migration Patterns: Experimental Evidence*

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

Abstract

We experimentally analyze the effect of the endogenous choice of sanctioning institutions on cooperation and migration patterns across societies. In our experiment, subjects are allocated to one of two groups, are endowed with group-specific preferences, and play a public goods game for 30 periods. Each period, subjects can move between groups and, at fixed intervals, can vote on whether to implement formal (centralized) sanctioning institutions in their group. We compare this environment to one in which only one group is exogenously endowed with sanctioning institutions. We find that subjects' ability to vote on institutions leads to (*i*) a more efficient partition of subjects between groups, (*ii*) a lower migration rate, (*iii*) an increase in overall payoffs, and (*iv*) a decrease in both inter- and intra-groups (payoff) inequality. Over time, subjects tend to vote for sanctioning institutions and contribute to the public good.

Keywords: Formal Sanctions, Cooperation, Migration, Voting, Experiment.

JEL Classification Numbers: C73, C91, C92, D72, H41, H73.

*We are particularly indebted to Gary Charness and Carolina Di Tella for very useful comments and discussions. We also thank seminar participants at the University of Exeter BID Brown Bag seminar. The authors are thankful to the University of Exeter BID and BIP Clusters for financially supporting this project. All remaining errors are ours.

[†]Department of Economics, University of Exeter. E-mail: R.Cobo@exeter.ac.uk.

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

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

1 Introduction

Institutions, defined as "the humanly devised constraints that shape human interaction" (North, 1990, p. 3), play a fundamental role in societies. The extensive literature on institutions finds that "good" rules (e.g., sanctioning institutions and, more generally, the rule of law) help societies solve social dilemmas and positively affect many outcomes, such as productive effort and investments (Ostrom, Walker and Gardner, 1992; Acemoglu, Johnson and Robinson, 2001). Different societies may, however, implement different rules and, as a consequence, experience contrasting development paths (North and Thomas, 1973; Jones, 1981; North, 1981; Acemoglu, Johnson and Robinson, 2005). These differences, in turn, can affect individuals' decisions to move between societies and, as observed in laboratory experiments, eventually generate strong migration patterns (Gürerk, Irlenbusch and Rockenbach, 2006).

Migration flows caused by differences in communities' ability to solve social dilemmas might entail costs for both individuals and societies. First, because of social and family ties, individuals can have preferences for their "home society" and would therefore bear a cost from moving away from home (Hill, 1987; Niedomysl and Amcoff, 2011). Second migration can negatively affect the "sending society" – and widen the welfare gap between this society and the "hosting" one – by, for instance, depleting part of its workforce (Katseli, Lucas and Xenogiani, 2006; Beine, Docquier and Rapoport, 2007; De Haas, 2010).[1]

Provided that heterogeneity in institutions is one of the factors driving migration flows and leading to inequalities between societies, analyzing whether and to what extent institutions themselves can be a solution to these issues then becomes important. In this paper, we use a laboratory experiment to investigate whether the possibility of individuals self-organizing by choosing the institutions governing their interactions regulates migration flows efficiently and mitigates inequality between societies.

[1] The cost of migration for the sending society grows larger if positive spillovers exist between workers, for instance in terms of productivity (see Katseli, Lucas and Xenogiani, 2006). Migration flows can also have (long-run) positive effects on the sending society (De Haas, 2010) as well as positive and negative effects on the hosting society (Friedberg, 2001; Dinkelman and Schulhofer-Wohl, 2015). In this paper, we abstract from these additional effects.

We focus on a particular set of institutions: sanctions. Recently, experimental economists have devoted considerable attention to the emergence of sanctioning institutions (Kosfeld, Okada and Riedl, 2009) and their impact on cooperation in public good games. Informal (peer-to-peer) and formal (centralized) punishment mechanisms have been shown to significantly increase the level of cooperation (Ostrom, Walker and Gardner, 1992; Fehr and Gächter, 2000, 2002; Gächter and Herrmann, 2009; Van der Heijden, Potters and Sefton, 2009; Gürerk, Irlenbusch and Rockenbach, 2009).[2] We take this finding as a starting point for our experimental design. In our baseline treatment, we employ a setting similar to Gürerk, Irlenbusch and Rockenbach (2006). At the beginning of the experiment, we randomly allocate subjects to two groups to play a Public Goods Game (PGG). Subjects interact for 30 periods. One of the groups implements sanctioning institutions, whereas the other group has no rules governing the interactions between its members. Every period, participants decide whether to move between groups. Three main features distinguish our baseline from Gürerk, Irlenbusch and Rockenbach (2006). First, we endow subjects with group-specific preferences; that is, all else equal, subjects obtain higher payoffs in the group to which they are initially allocated ("initial group"). Second, to account for the costs endured by a society as a result of migration outflows, we make the individual return from the public good depend on the group size. Third, we use formal rather than informal sanctioning institutions.

We compare this baseline with our main treatment, in which subjects also have the possibility of voting at fixed intervals on the institutions – either formal sanctions or no sanctions – to be implemented in their group. We focus on endogenous institutional choice over exogenously imposed institutions mainly because (i) imposing institutions is not always possible in democratic societies, and (ii) institutions are more effective when chosen by individuals rather than exogenously implemented (Sutter, Haigner and Kocher, 2010; Baldassarri and Grossman, 2011; Markussen, Putterman and Tyran, 2014).

Although the possibility of migrating and the interpretation of sanctioning institutions

[2]For a survey of this literature, see Chaudhuri (2011).

as a bundle of local taxes and public goods link our paper to Tiebout (1956), our research question differs from that answered by Tiebout-like models. Rather than studying the effect of "voting with feet" on individuals' optimal sorting given preferences over taxation and public goods provision, we study the effect of the endogenous choice of institutions on individuals' sorting given preferences for location.

We find that the endogenous institutional choice dramatically affects migration, individual welfare, and inequality patterns. Unlike the case in which individuals cannot vote on institutions, groups that can choose the rules governing their interactions do not significantly diverge over time in terms of their size, composition, and contributions. On average, over the 30 periods of the game, less than 73% of subjects play in their initial group when voting is not allowed, with the group endowed with sanctioning institutions attracting most of the subjects. By contrast, when institutional choice is allowed, a statistically significant higher share (83%) of subjects play in their initial group, with the majority (62%) voting to implement sanctions by the end of the experiment. Group sizes are then more balanced when subjects can vote on institutions than in the case in which institutions cannot be chosen. Results also show that although contributions are very similar with and without institutional choice, the possibility of self-organizing leads to higher payoffs for individuals belonging to an otherwise depopulated society, and a more equal distribution of payoffs between groups. The individual-level analysis confirms these findings and points out subjects' propensity to move back to their initial group to cast their ballot, which further emphasizes their strong ability to self-organize efficiently.

The paper proceeds as follows. Section 2 presents a review of the literature. Section 3 describes the experimental design. Section 4 introduces a formal model and presents the main theoretical predictions and the hypotheses to be tested in the empirical analysis. Section 5 reports and discusses our findings. Section 6 concludes.

2 Literature Review

Our paper continues a stream of literature that studies both the effect of institutions on human cooperation and the evolutionary properties of sanctioning institutions.

Since the work of Yamagishi (1986) and Ostrom, Walker and Gardner (1992), many scholars have investigated the role played by the possibility of punishing free-riders in fostering cooperation in social dilemmas. Fehr and Gächter (2000, 2002) show individuals engage in decentralized altruistic punishment – that is, they punish defectors at a cost to themselves and in the absence of material gains – which leads to high levels of cooperation. Baldassarri and Grossman (2011) study the effectiveness of centralized forms of punishment (e.g., a police force) in sustaining cooperation by using a lab in the field experiment. The authors show centralized sanctions enhance cooperation, especially when the centralized authority is elected rather than randomly chosen.

Recently, increasing attention has been devoted to the endogenous choice of different types of decentralized institutions. Ertan, Page and Putterman (2009) study the adoption of different punishment technologies. The authors find individuals tend to increasingly implement sanctioning institutions over time, and choose to punish only low contributors to the public good. Sutter, Haigner and Kocher (2010) study the endogenous choice of sanctions and rewards within a standard Voluntary Contribution Mechanism. They find individuals choose the rewarding or punishment options, as opposed to a sanction-free environment, when sufficiently effective; also, individuals tend to prefer to reward cooperation rather than punish free-riding. An effect of the endogenous choice of institutions is found: higher cooperation levels arise when individuals endogenously choose the rules governing their interactions.

Because the presence of second-order free-riding and antisocial punishment lowers the effectiveness of peer-to-peer punishment (Denant-Boemont, Masclet and Noussair, 2007; Dreber et al., 2008; Herrmann, Thöni and Gächter, 2008; Nikiforakis, 2008), a strand of this literature focuses on the emergence of centralized forms of punishment. Andreoni and Gee (2012) experimentally compare decentralized to centralized punishment and find that in the vast majority of cases, subjects implement a centralized mechanism that crowds out the peer-to-peer mechanism. Sigmund et al. (2010) theoretically show more centralized forms of punishment (i.e., pool punishment) are more stable than peer punishment, because second-order free-riding is also punished. The authors, however, find pool punishment to be less efficient than peer punishment. Zhang et al. (2014) generally confirm this prediction by means of an experiment, even though they find peer punishment to be more stable than theoretically predicted. Traulsen, Röhl and Milinski (2012) argue a centralized punishment is preferred to a decentralized one provided individuals are punished for not contributing to the funding of centralized institutions.

Markussen, Putterman and Tyran (2014) also focus on individuals' dynamic choice of punishment mechanisms. The authors experimentally analyze the conditions under which a centralized punishment mechanism emerges as the preferred mechanism vis-à-vis an informal mechanism as well as a sanction-free environment. They find that, in line with social preferences theories, a decentralized punishment mechanism is often preferred to a centralized mechanism whenever the latter becomes sufficiently costly (but profitable). By contrast, a centralized mechanism becomes the preferred choice when deterrent and sufficiently cheap.[3] As in Sutter, Haigner and Kocher (2010), the authors further confirm the welfare-enhancing effect of the endogenous institutional choice.

Gross et al. (2016) analyze individuals' decisions to delegate punishment power to others. The authors show hierarchical power structures emerge out of voluntary transfer of punishment power to those members who are more willing to engage in decentralized punishment; this process results in sustainable cooperation.

Our paper complements this literature by studying the evolution and regulatory properties of sanctioning institutions when subjects can vote with both feet and ballots.

Subjects' ability to move between groups also links our paper to the literature on the relationship between institutions and migration patterns. Gürerk, Irlenbusch and Rockenbach

[3]On the advantage of centralized forms of punishment over decentralized ones, see also Guala (2012).

(2006) experimentally analyze a public goods game and show that, when confronted with a sanction-free society, a society endowed with peer-to-peer punishment succeeds in attracting all the subjects and sustaining cooperation. In a similar environment, Gürerk, Irlenbusch and Rockenbach (2014) find individuals' self-selection into the community implementing the preferred institutional framework is a key determinant of long-run cooperation. We mainly differ from Gürerk, Irlenbusch and Rockenbach (2006) in that we allow each group to endogenously choose its preferred sanctioning regime.

Robbett (2014) studies a Tiebout-like environment in which (*i*) individuals are endowed with preferences over bundles of local taxes and public good provision, and (*ii*) returns from the public good increase with community size. The author finds that, when unable to vote on their preferred bundle, individuals tend to sort optimally but fail to achieve full efficiency due to over-segregation in communities providing a Pareto-dominated bundle; this inefficiency is overcome when individuals can vote on bundles. Our research question is related to but differs from that answered by Tiebout-like models because, as noted before, we focus on the effect of the endogenous choice of institutions on individuals' sorting given preferences for location.

3 Experimental Design and Procedures

We conducted laboratory experiments at the University of Exeter between February and October 2016. Participants were mainly students of economics, business administration, and engineering, but other disciplines were also represented. We ran two different treatments with a total of 50 sessions, 25 sessions per treatment. Each session was composed of 10 subjects; the average individual earnings were £14. All instructions can be found in Appendix A.

The experiment comprises two different activities. Upon arrival to the laboratory, subjects are randomly divided into two groups – A and B – of five people each. In the first activity, each subject receives 50 tokens and participates in a standard one-shot PGG. At the end of the first activity, subjects receive the instructions for the second task, which consists of a *PGG* that lasts 30 periods. No information about the outcome of the first activity is released. Treatments for the second activity differ depending on whether the institutional setting is endogenously chosen.

In **Treatment 1** (*No-Voting – Moving*), subjects are randomly assigned in the first period to one of two groups – A and B – of five people each. The group to which each subject is allocated in the first period is her "initial group". In each period, each subject is endowed with 50 tokens to be allocated between a group and a private account. As Table 1 shows, the subject's return from an allocation to the group account depends on the size of the group the subject is located in, and on whether this group is the subject's initial group.

Table 1: Individual returns to the public good by group size and membership

Group Size (n)	**Factor If Subject Is In Her Initial Group**	**Factor If Subject Is *Not* In Her Initial Group**
1	1.00	1.00
2	1.50	1.15
3	1.75	1.40
4	2.00	1.65
5	2.25	1.90
6	2.25	1.90
7	2.25	1.90
8	2.25	1.90
9	2.25	1.90
10	2.25	1.90

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

(a) Under Rule-Set 1, sanctioning institutions are not in place and individual i's per-period monetary payoff is

$$\pi_{i,h} = (50 - C_i) + \frac{R_h(n)}{n} \sum_{j=1}^{n} C_j \tag{1}$$

for $i, j = 1, ..., n$, where n denotes the total number of individuals located in the group, $C_i \in [0, 50]$ denotes i's contribution to the group account, and $\frac{R_h(n)}{n}$ represents the marginal per-capita return ($MPCR$) from the group account for individual i whose initial group is $h = A, B$, with values for $R_h(n)$ shown in Table 1.

(b) Under Rule-Set 2, sanctions are in place. Each individual in the group has to pay a fixed fee of 15 tokens per period. In addition, each individual pays a fine equal to 80% of the amount of tokens allocated to the private account in a particular period. The fixed fee and the fine (if applicable) are deducted from subjects' monetary payoff at the end of the period. Under Rule-Set 2, i's per-period monetary payoff is therefore

$$\pi_{i,h} = (50 - C_i)(1 - 0.8) + \frac{R_h(n)}{n}(\sum_{j=1}^{n} C_j) - 15. \quad (2)$$

In Treatment 1, Group A implements Rule-Set 1 and Group B implements Rule-Set 2 for the entire duration of the game.

At the end of each period, subjects receive information about (i) the average contribution to the group account in their current group, (ii) the average contribution in the other group, (iii) the rule-set implemented in their current group, and (iv) the rule-set implemented in the other group. Subjects then decide whether to move from their current group.

In **Treatment 2** (*Voting – Moving*), besides deciding the contribution to the group account and whether to move between groups, subjects also vote on the rule-set to be implemented.

In the first five periods, Group A implements Rule-Set 1, and Group B implements Rule-Set 2. Starting from period $t = 6$, subjects vote every 5 periods on the rule-set to be implemented in their current group (i.e., the group in which they play in that particular period, which may differ from their initial group). Subjects therefore vote five times in total throughout a session. Information about the rule-set chosen by the majority of voters is publicly released to all subjects; the rule-set is implemented immediately after voting and applies until the next voting round.[4]

[4]In case the two rule-sets receive the same number of votes, the rule-set for the group is randomly chosen.

4 Setup

In this section, we present the model setup and derive the theoretical predictions arising from our setting. We consider two groups composed of n^g individuals, with $g = A, B$, denoting the group index, and where $N = n^A + n^B$ represents the total number of individuals in our economy. Individuals play a PGG for a finite number of periods T.

In each period $t \in [1, T]$, players first simultaneously choose their contribution to the production of the public good, and then simultaneously decide whether to move between groups. Contributions can occur under two regimes. In the *sanction-free* regime, individual i's contribution $C_i \in [0, E]$ is voluntary, where E denotes i's endowment, for $i = 1, ..., N$. In the *sanctioning institutions* regime, i's contribution is subject to a fine equal to a fraction $s \in (0, 1]$ of the endowment she does not allocate to the public good.

Unlike in Treatment 1, individuals in Treatment 2 vote every k periods on whether to implement formal sanctions in their group (g). Voting occurs at the beginning of the period, and contributions are chosen once the outcome of majority voting is implemented and made publicly known.

Sanctioning institutions and contributions to the public good are local, and no spillovers exist between groups. The cost of sanctioning institutions is proportional to the size of the group: each individual playing under sanctioning institutions must contribute a fixed amount c to its provision (e.g., to fund law courts and police forces).

Let h define individuals' "initial group"; that is, the group i is allocated to in $t = 1$, for $h = A, B$. We assume each group has $n_h = n^*$ "initial group members", with $n_A + n_B = N$.[5] We denote by $\tilde{n}^g \in [0, n^*]$ the number of members in g who are also initial group members, for $g = A, B$.

In the absence of sanctions, the per-period monetary payoff of individual i, with initial

[5]Whenever h and g are employed, we adopt the convention of reporting individuals' location as a superscript and individuals' initial group as a subscript.

group h and located in group g of size n^g, is

$$\pi^g_{i,h} = (E - C^g_{i,h}) + \frac{R^g_{i,h}(n^g)}{n^g} \sum_{j=1}^{n^g} C^g_j \tag{3}$$

for $i, j = 1, ..., n^g$, and $h, g = A, B$. In (3), $r^g_{i,h}(n^g) \equiv \frac{R^g_{i,h}(n^g)}{n^g}$ represents the $MPCR$ from the public good, where $R^g_{i,h}(\cdot)$ multiplies i's share of the public good to capture the social utility of the public good provision, for $h, g = A, B$. We assume the following:

A1: $R^g_{i,h}(n^g)$ is non-decreasing in n^g and attains its maximum at $n^g = n^*$, for $h, g = A, B$.

A2: $1/n^g \leq r^g_{i,h}(n^g) \leq 1$ for $i = 1, ..., n^g$, and $h, g = A, B$.

A3: $R^h_{i,h}(n) > R^g_{i,h}(n)$ (and therefore $r^h_{i,h}(n) > r^g_{i,h}(n)$), where $R^h_{i,h}(1) < R^g_{i,h}(N)$, for $h, g = A, B$ and $h \neq g$, and for $n = 1, ..., N$.

From **A1**, the social benefit from the provision of the public good – relative to its cost – increases with the group size up to n^* and remains constant for $n^g \in (n^*, N]$. **A1** may, for instance, reflect economies of scale in the production of the public good (e.g., education). **A2** gives rise to a social dilemma: in (3), (self-interested) i has an incentive to free-ride and contribute $C^g_{i,h} = 0$, whereas efficiency requires $C^g_{i,h} = E$, for $h, g = A, B$. **A3** captures i's disutility from not residing in h. However, the last inequality in **A3** implies the "initial group premium" is not sufficiently high to compensate i for living in a largely depopulated h. Together, **A1-A3** imply a *socially efficient partition of individuals* between groups: for any given level of (group) contribution, the public good is most beneficial to a society when each individual resides in her initial group; that is, $n^g = \tilde{n}^g = n^*$, for $g = A, B$.

When g adopts sanctions, the per-period monetary payoff of individual i, with initial group h and located in g, is given by

$$\pi^g_{i,h}(s) = (E - C^g_{i,h})(1 - s) + \frac{R^g_{i,h}(n^g)}{n^g} \left(\sum_{j=1}^{n^g} C^g_j \right) - c. \tag{4}$$

We assume sanctions are *deterrent*; that is, they make it individually rational to contribute the entire endowment to the public good:

A4: $s > (1 - r^{g}_{i,h}(n^{g}))$, $\forall n^{g}$, for $h, g = A, B$.[6]

Following Markussen, Putterman and Tyran (2014), we define $\mathcal{P}^{g}_{i,h}(n^{g}) = R^{g}_{i,h}(n^{g})E - E$ as the "cooperation premium", that is, the increment in the earnings of i with initial group h and located in g due to the presence of formal sanctions.[7] The cooperation premium is individual-specific: it increases as (a) an individual joins her initial group, and (b) the group size rises toward the efficient size n^{*}. We assume:

A5: $c < \mathcal{P}^{g}_{i,h}(n^{g})$ if $n^{g} > \underline{n}^{g}_{h}$, for $h, g = A, B$, where $n^{*} > \underline{n}^{g}_{h} > \underline{n}^{h}_{h}$, for $g \neq h$.

From **A5**, when individuals free-ride absent formal sanctions, implementing a formal-sanction scheme generates a net benefit for an individual provided the group size is larger than the threshold $\underline{n}^{g}_{h}$. Because the sanctions' cooperation premium increases when the individual resides in her initial group, we have $\underline{n}^{g}_{h} > \underline{n}^{h}_{h}$, for $g \neq h$.

4.1 Predictions under Selfish Preferences

We now derive the equilibrium predictions arising from our setting, assuming it is common knowledge that all individuals are self-interested and rational. We first discuss individual contribution, migration and, for Treatment 2, voting behavior. We then present equilibrium predictions for Treatment 1 (which offers no possibility of voting) and Treatment 2 (in which subjects can vote on sanctioning institutions).

[6] From Table 1, **A4** is marginally violated for individuals not located in their initial group when $n^{g} = N$, for $g = A, B$ – the difference between the left-hand side and the right-hand side in **A4** being 0.01. This discrepancy affects only one equilibrium prediction (see point (a) in the *Strong Nash Stable Equilibrium* described below) in terms of the contribution of individuals not located in their initial group (because it predicts a zero contribution), but not in terms of location and voting behavior. However, we find no evidence of sanctions being non-deterrent in the case in which subjects are all located in one group. For the sake of simplicity, we therefore present a theoretical analysis based on **A4**.

[7] We define earnings as the individual monetary payoff net of the cost of formal sanctions. Notice our definition of $\mathcal{P}$ assumes individuals are self-interested and therefore do not contribute to the public good absent formal sanctions.

4.1.1 Contribution

From **A2**, absent sanctioning institutions, individual i contributes $C^g_{i,h} = 0$, for $i = 1, ..., n^g$, and $h, g = A, B$. Also, from **A4**, individual i contributes $C^g_{i,h} = E$ in the presence of sanctions, for $i = 1, ..., n^g$, and $h, g = A, B$.

4.1.2 Migration

Let n^g_t denote the size of g in period t. Also, let $m^g_{i,h,t}$ be an indicator taking the value 1 if individual i, whose initial group is h, decides to move from her current group g in period t, and 0 otherwise, for $h, g = A, B$ and $i = 1, ..., n^g_t$.

Remark 1. Absent sanctioning institutions in both groups, an individual is indifferent between her "initial group" and the "non-initial group".

Remark 2. Suppose only group B is endowed with formal sanctions in $t + 1$. If i, for $i = 1, ..., n^A_t$, is located in her initial group A and believes $n^B_{t+1} > \underline{n}^B_A$ (resp., $n^B_{t+1} \leq \underline{n}^B_A$), then $m^A_{i,A,t} = 1$ (resp., $m^A_{i,A,t} = 0$). Similarly, if i, for $i = 1, ..., n^B_t$, is located in her initial group B and believes $n^B_{t+1} \leq \underline{n}^B_B$ (resp., $n^B_{t+1} > \underline{n}^B_B$), then $m^B_{i,B,t} = 1$ (resp., $m^B_{i,B,t} = 0$).

Remark 3. If both groups implement sanctioning institutions, individuals obtain the highest payoff when located in their initial group.

By migrating, individuals express preferences for the public good provision and/or groups. "Voting with feet" is, however, powerless absent sanctions, because neither of the two groups contributes to the public good (Remark 1).

When only one group implements sanctioning institutions, different beliefs support different migration patterns. In particular, from Remark 2, because the sanctions' cooperation premium ($\mathcal{P}^g_{i,h}$) increases with the group size, subjects may find it individually rational to move to the group in which all the other subjects are located, regardless of the sanctioning regime implemented.

This statement also holds when both groups implement formal sanctions (Remark 3). In this case, however, each subject can find it individually rational to settle in her initial group if other subjects do so as well.

4.1.3 Voting

From **A5**, a threshold $\underline{n}_h^g$ exists for $h, g = A, B$, such that voting for sanctions is a dominant strategy if $n^g > \underline{n}_h^g$, where $\underline{n}_h^g > \underline{n}_h^h$, for $g \neq h$.

Remark 4. All else equal, when playing in her initial group, an individual enjoys a higher cooperation premium from sanctioning institutions. Therefore, individuals are more likely to vote in favor of sanctioning institutions when located in their initial group.

4.1.4 Equilibrium Outcomes

Treatment 1. When voting on institutions is not allowed and only Group B is endowed with sanctioning institutions, the following equilibria exist.

Nash Equilibria with Exogenous Institutions. *In Treatment 1, two equilibria exist in which all individuals are located in either Group A or Group B. Individuals contribute to the public good only when located in the group implementing sanctions (Group B). Hence, the equilibrium in which all individuals are located in Group B is Pareto-dominant.*

Because we expect the Pareto-dominant equilibrium to arise, we have the following hypothesis:

Hypothesis 1. *In Treatment 1, we expect:*

(***1.1***) *Group B to become overpopulated relative to Group A by attracting A's initial members;*

(***1.2***) *subjects to contribute more when sanctioning institutions are in place;*

(***1.3***) *payoff inequality to emerge both (i) between groups and (ii) within the overpopulated group (Group B), between its initial and non-initial members.*

Treatment 2. In Treatment 2, only Group B implements sanctioning institutions in $t = 1, ..., k_1 - 1$, where $k_1 < T$ denotes the first voting round. Voting occurs at regular intervals of k periods: the first voting round occurs at the beginning of period $k_1 \equiv k + 1$. Following Robbett (2014), we focus on *Nash stable* and *strong Nash stable* partitions of individuals.[8] A partition is *Nash stable* if no individual can obtain a higher payoff from unilaterally moving to a different group. A partition is *strong Nash stable* if no coalition of agents exists in which all weakly prefer *coalitionally* moving to a different group, with at least one member of the coalition being strictly better off.

Nash Stable Equilibria. *In Treatment 2, two Nash Stable equilibria exist in which all individuals are located in either Group A or Group B in $t = 2$, no migration occurs in the remaining periods, all individuals vote in favor of formal sanctions in every voting round, and individuals contribute to the public good only when located in the group implementing sanctions.*

Initial migration driven by beliefs and, possibly, different institutional settings between groups may lead to inefficient outcomes. When all subjects are located in one group, none of them finds it individually rational to migrate further (Remark 2). The two groups then diverge in terms of welfare, as one group becomes fully depopulated. Also, payoff inequality

[8] On equilibrium selection in Tiebout models, see Greenberg and Weber (1986) and Conley and Konishi (2002).

within the populated group arises as the initial group members obtain higher returns from the public good.

Strong Nash Stable Equilibrium. *In Treatment 2, a Strong Nash Stable equilibrium exists in which*

(a) individuals contribute 0 *(resp.,* E*) when located in a sanction-free group (resp., in a group implementing sanctions);*

(b) all individuals are located in Group B in $t = 2, ..., k_1 - 1$*;*

(c) all individuals whose initial group is A move to Group A at the end of period $t = k_1 - 1$*;*

(d) in $t = k_1$*, all individuals vote in favor of formal sanctions;*

(e) no migration occurs from $t = k_1$ *onward, and all individuals vote in favor of sanctioning institutions in all subsequent voting rounds.*

The *Strong Nash Stable Equilibrium* describes the Pareto-dominant outcome that is made possible by endogenous institutional choice. Subjects initially move to the group implementing sanctioning institutions to enjoy the efficient provision of the public good. Right before the first voting round – that is, in the "pre-voting" period – individuals located outside their initial group coalitionally move and implement sanctioning institutions in their initial group. Once both groups are endowed with sanctioning institutions, this efficient partition of individuals becomes the unique strong Nash stable partition. This equilibrium also eradicates payoff inequalities both between and within groups.[9]

Because the Strong Nash Stable Equilibrium is Pareto-dominant, we expect this equilibrium to arise. Therefore, we have the following hypothesis:

[9] Because groups are populated by initial members only, there is no inequality within groups.

Hypothesis 2. *In Treatment 2, we expect:*

(***2.1***) *groups to converge toward the efficient partition of individuals; that is, the majority of subjects are located in their initial group;*

(***2.2***) *the majority of subjects to vote for sanctioning institutions;*

(***2.3***) *subjects to contribute more when sanctioning institutions are in place;*

(***2.4***) *no payoff inequality between groups;*

(***2.5***) *subjects located outside their initial group to be more likely to move back in the "pre-voting" period (vis-à-vis other periods).*

4.2 Predictions under Social Preferences

We briefly present some of the consequences of assigning social preferences – e.g., preferences for reciprocity and/or efficiency (Fehr and Schmidt, 1999; Charness and Rabin, 2002) – to individuals for equilibrium dynamics, as a formal analysis proves particularly cumbersome in the presence of migration.[10]

In Treatment 2, social preferences can give rise to the following equilibrium prediction.

Equilibrium under Social Preferences. *Suppose social preferences stabilize cooperation at a sufficiently high level in both groups absent sanctioning institutions. Then, in Treatment 2, an equilibrium exists in which individuals are efficiently located in their sanction-free initial groups.*

This equilibrium describes a particularly efficient result: individuals contribute to the public good in their initial group without incurring the cost of sanctioning institutions. This equilibrium can arise in the presence of a sufficiently strong aversion to advantageous inequality (Fehr and Schmidt, 1999) and/or a sufficiently strong preference for social welfare and overall efficiency (Charness and Rabin, 2002).

[10]For a formal discussion of the impact of social preferences on individuals' voting behavior on formal sanctioning institutions in the absence of migration, see Markussen, Putterman and Tyran (2014).

5 Results

We start by comparing the dynamic and convergence results of the game attained when subjects can only move between groups (*No-Voting-Moving*) to those attained when subjects can decide on both the rule-sets and their location (*Voting-Moving*). We then examine the main determinants of subjects' contribution, migration, and voting decisions.

5.1 Dynamics and Convergence

Dynamics. We start with an overview of groups' dynamics in the two treatments. The top (bottom) panel in Figure 1 plots the average size of Groups A and B over time in the *No-Voting-Moving* (*Voting-Moving*) treatment. The figure further discriminates between subjects whose current group is/is not their initial group.

Similar to Gürerk, Irlenbusch and Rockenbach (2006), in the *No-Voting-Moving* treatment (*NVM* henceforth), Group A – in which no sanctioning institutions are in place – becomes depopulated over time, whereas Group B – in which sanctions are in place – becomes overpopulated. As the top panels of Figure 1 show, the size of Group A goes from five members in the initial period to about two members on average in the last five periods of the experiment.[11] At the same time, Group B increases from five members in the initial period to almost eight members on average in the last five periods.[12] As a result, the average size of Group A is significantly lower than that of Group B.[13]

The upper panels in Figure 1 also reveal the composition of both groups varies throughout the experiment. In particular, Group B increases in size by attracting individuals who were initially allocated to Group A. In the first five periods, only 18.7% of the subjects initially allocated to Group A (0.94 subjects on average) are located in Group B. By the last five periods, that percentage grows to 59.4% (2.97 subjects), a statistically significant increase

[11] Differences are statistically significant (Wilcoxon signed rank test: z=4.20, p=0.00, two-tailed).
[12] Differences are statistically significant (z=-4.19, p=0.00, two-tailed).
[13] The average size of Group A over the 30 periods is 3.15 across all sessions, while it is 6.85 for Group B. Differences are statistically significant (Wilcoxon signed rank test: z=-4.37, p=0.00, two-tailed).

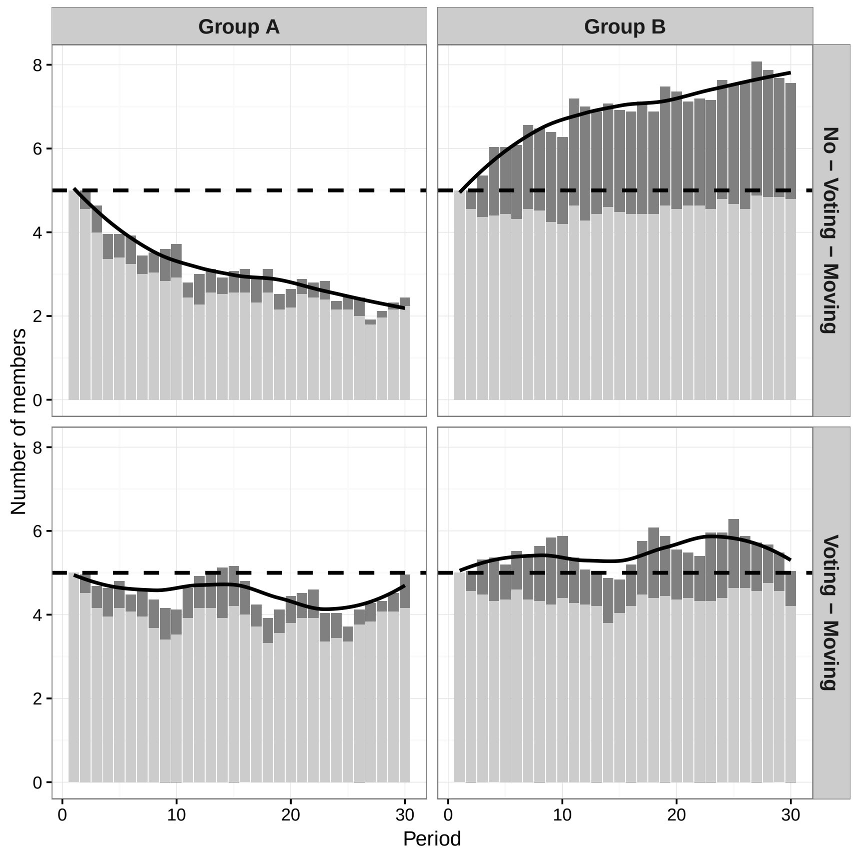

Figure 1: Distribution of participants across groups over time, by treatment. Lighter (darker) portions of the vertical bars represent the average number of individuals for whom their current group is (is not) their initial group. Solid lines represent locally weighted scatterplot smoothing curves fitted to the data. The dashed horizontal lines represent the situation in which participants are equally distributed between Group A and Group B.

(Wilcoxon signed rank test: z=4.21, p=0.00, two-tailed). These results support Hypothesis 1.1.

The bottom panels in Figure 1 show a more equal distribution of subjects between Group A and Group B in the *Voting-Moving* treatment (VM henceforth) than in NVM. The two treatments differ both in the size and composition of the groups. First, in VM, the average number of members in both Group A and Group B is close to five, and remains quite stable over the course of the experiment. The average number of members in each group in the first five periods is not significantly different from that in the last five periods (Wilcoxon signed rank test: z=1.20, p=0.24 for Group A, and z=-1.16, p=0.25 for Group B, two-tailed). Second, most of the members in Group A and Group B play in their initial groups. On average, 78% (88%) of subjects whose initial group is Group A (B) play in Group A (B), and these proportions do not vary significantly over the 30 periods of the experiment: the p-value for a permutation test for repeated measures (Pesarin and Salmaso, 2010) is 0.11. Altogether, 83% of subjects in *VM* are located in their initial group in any given period, in accordance with Hypothesis 2.1. The corresponding percentage for NVM (72.7%) is significantly lower (Mann-Whitney test: z=-3.23, p=0.00, two-tailed).

To better understand the differences in the size and composition of groups between NVM and *VM*, Figure 2 looks at subjects' migration behavior. The left panel plots the average share of subjects moving between groups in each treatment; the right panel distinguishes moves toward groups that implement sanctions from moves toward sanction-free groups.

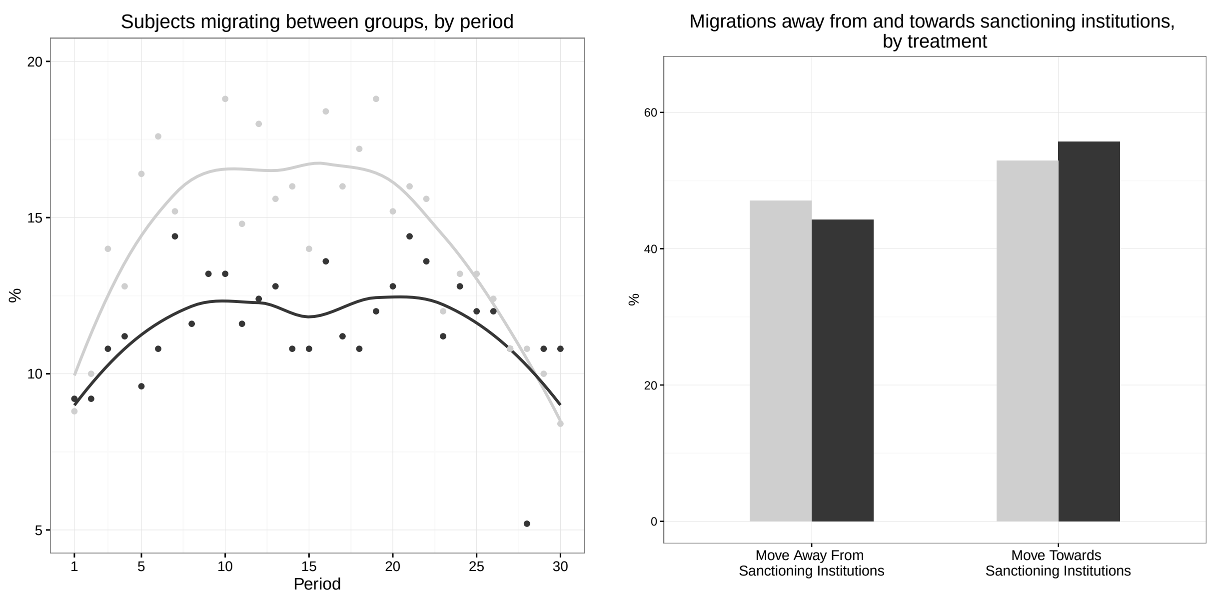

Figure 2: Migration patterns under the *Voting-Moving* and *No-Voting-Moving* treatments. The left panel plots the average percentage of moves between groups over time. Black (gray) dots represent the proportion of moves in each period for *VM* (NVM); solid lines represent locally weighted scatterplot smoothing curves fitted to the data. The right panel plots the proportion of moves away from/toward groups implementing sanctions – out of the total number of group shifts – under *VM* (black bars) and *NVM* (gray bars). For *VM*, the sample in the right panel comprises only periods-sessions in which the rule-sets implemented by Group A and Group B differ.

As shown in the left panel, the percentage of subjects who move between groups over the course of the experiment in *NVM* (14.76%) is larger than in *VM* (11.52%), and this difference is statistically significant (Mann-Whitney test: z=1.97, $p<0.05$, two-tailed).[14]

The right panel shows subjects are more likely to move toward groups that implement sanctions. Considering only those periods/sessions in *VM* in which the two groups implement different rule-sets (ca. 41% of the observations), we find 55.7% of the movements go from sanction-free groups to groups implementing sanctions, while 44.3% of the movements go in the opposite direction. In *NVM*, the corresponding figures are 52.9% and 47.1%, respectively. In both treatments, the proportion of movements toward sanctioning institutions is significantly higher than the proportion of movements away from institutions (exact binomial test: p=0.02 for *VM*, p=0.05 for *NVM*, two-tailed).

Overall, whereas the right panel of Figure 2 points to quite similar migration choices in *NVM* and *VM*, Figure 1 shows the dynamics in terms of groups' size and composition differ substantially between treatments. We attribute this difference to the ability of subjects to vote for their preferred institutions in *VM*.

Next, we analyze subjects' contribution to the public good over time, because this analysis can help account for subjects' propensity to move toward groups endowed with sanctions. As Figure 3 shows, in both *NVM* and *VM*, the average contribution in sanction-free groups is significantly lower than in groups implementing sanctions. In *NVM*, the average contributions in groups with and without sanctions are 45.67 and 20.55 tokens, respectively; in *VM*, the corresponding values are 46.61 and 24.75. These differences are statistically significant (Wilcoxon signed rank test: z=4.37, p=0.00 for *NVM* and z=4.37, p=0.00 for *VM*, two-tailed), consistent with Hypotheses 1.2 and 2.3.

Finally, Figure 4 analyzes subjects' average payoffs over time in the two treatments. In *NVM*, the average per-period payoffs are 67.30 for Group A members and 82.17 for Group B members, a statistically significant difference (Wilcoxon signed rank test: z=4.34, p=0.00,

[14]We also compared the proportion of movers under *NVM* and *VM* using a permutation test, which imposes fewer assumptions and has higher power than the Mann-Whitney test (Pesarin and Salmaso, 2010). The null hypothesis of no difference between treatments is again rejected (p-value=0.01, two-tailed).

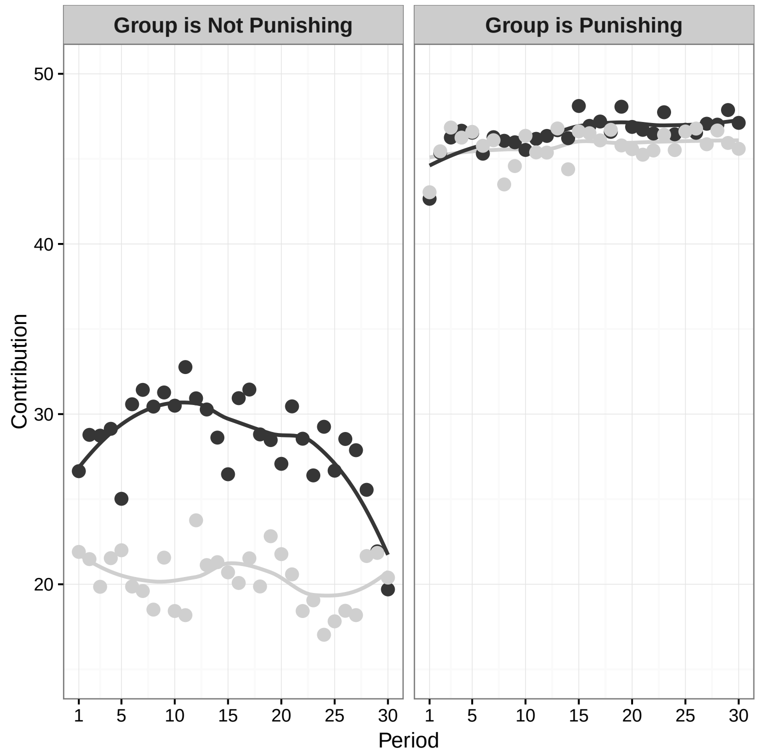

Figure 3: Average group contribution per period, by institutional setting and treatment. The left (right) panel plots the average contribution for groups without (with) sanctioning institutions. In both cases, black (gray) dots represent the mean contribution per period under the *VM* (*NVM*) treatment; solid lines represent locally weighted scatterplot smoothing curves fitted to the data.

two-tailed). Furthermore, among subjects located in Group B, the average payoff of initial Group B members (87.84) is significantly higher than that of non-initial members (72.41) (Wilcoxon signed rank test: z=4.37, p=0.00, two-tailed). Both results support Hypothesis 1.3.

In *VM*, the average per-period payoffs for Group A and Group B members are 81.18 and 85.15, respectively. Even though this difference is also significant (Wilcoxon signed rank test: z=2.27, p=0.02, two-tailed), contradicting Hypothesis 2.4, the disparity in payoffs between groups is significantly lower than in *NVM* (Mann-Whitney test: z=-4.26, p=0.00, two-tailed), suggesting that subjects' ability to vote on sanctioning institutions leads to more equal societies. Moreover, the average payoff across both groups in *VM* (84.07) is significantly higher than in *NVM* (78.47) (Mann-Whitney test: z=2.80, p=0.00, two-tailed), indicating

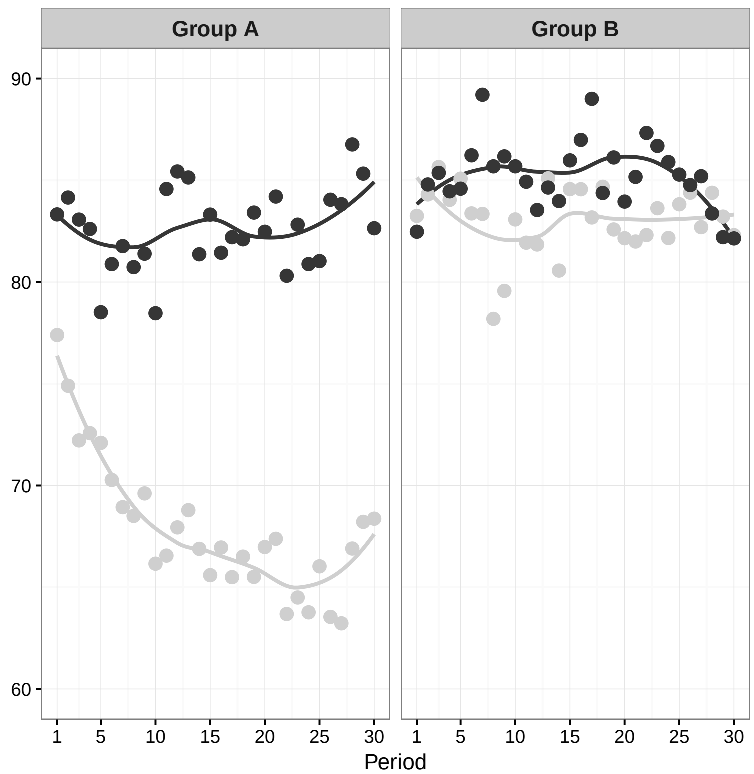

Figure 4: Average participant payoff per period, by current group and treatment. Black (gray) dots represent the average payoff per period of those subjects located in Groups A and B in each period, under the VM (NVM) treatment; solid lines represent locally weighted scatterplot smoothing curves fitted to the data.

the ability to choose institutions increases overall efficiency.

In the following subsection we analyze the stability of our results by investigating groups' convergence.

Convergence. We focus on two main dimensions of convergence. We define "convergence in groups in period γ" as the situation in which, given a particular composition of groups in a given session, at most one subject moves between groups each period from γ onward. We also define "convergence in institutions in period τ" as the situation in which neither group changes its institutional setting from τ onward in a given session. Finally, we define "double convergence in periods γ and τ" as the situation in which convergence in groups is achieved by period γ and convergence in institutions is achieved by period τ.

The upper part of Table 2 shows, for each treatment, the percentage of sessions achieving group convergence at least three periods before the end of the experiment (i.e., by period $\gamma = 27$ or earlier), the average time of convergence, and the average percentage of subjects who play in their initial group from the convergence period onwards.[15]

Table 2: Convergence in groups and convergence in institutions

	No-Voting-Moving	Voting-Moving
Convergence in Groups ($\gamma = 27$)		
Converging Sessions (%)	52.00	48.00
Average Convergence Time	22.92	19.75
Subjects in Initial Group (%)	68.97	96.39
Convergence in Institutions ($\tau = 22$)		
Converging Sessions (%)		68.00
Average Convergence Time		11.00
“Two Groups with Sanctions" (%)		52.00
“Two Groups without Sanctions" (%)		24.00
“One Group with/without Sanctions" (%)		24.00

In *NVM*, 52% of the sessions converge in groups by period 27. On average, sessions that converge in groups do so in ca. 23 periods. In line with the predictions of the *Nash Equilibria with Exogenous Institutions*, the groups' composition is far from efficiency (68.97% of subjects play in their initial group). In *VM*, we observe similar results to those in *NVM* in terms of the percentage of converging sessions (48%) and their convergence time (19.75 periods). However, when convergence occurs, the two groups are almost entirely populated by their initial members (96.39% of subjects play in their initial group). In line with the predictions of the *Strong Nash Stable Equilibrium*, this finding points to subjects' ability to converge toward the efficient partition of individuals between groups in *VM*.

The lower part of Table 2 displays the results for the convergence in institutions. For this type of convergence, we report the results for *VM* only, because subjects cannot vote in *NVM*. The table shows that almost 70% of the sessions converge in institutions by period 22

[15]The main conclusions drawn from Table 2 remain broadly similar for alternative values of γ.

(the next-to-last voting period) and that the average convergence time is 11 periods.

The table also reports the percentage of sessions that converge to each possible combination of institutions in the two groups. In the majority of the cases (52%) in which there is convergence in institutions by period 22, both groups converge to implementing sanctions. This result is in line with the predictions of the *Strong Nash Stable Equilibrium* under selfish preferences and, in particular, supports Hypothesis 2.2. However, we also observe 24% of sessions converging to both groups not implementing sanctions. This observation is consistent with the *equilibrium under social preferences* described in Section 4.2.

Finally, 24% of the sessions converge to two groups implementing different institutions (i.e., only one group implements sanctions). These sessions do not behave in accordance with the equilibria described in Sections 4.1 and 4.2.

5.2 Individual Behavior

To better understand the mechanisms underlying the group-level patterns reported above, this section presents individual-level analyses of contribution, migration, and voting decisions, focusing primarily on our main treatment (*VM*).[16]

Contribution. The first column of Table 3 reports the marginal effects from a (doubly) censored regression model (Greene, 2012) in which the dependent variable is the individual contribution of subject i in period t.[17] We have the following explanatory variables: $Is\ Punishing_{g,t}$, a binary covariate that equals 1 if a sanctioning mechanism is in place in i's current group g, and 0 otherwise; $Initial\ Group_{i,t}$, a dummy taking the value 1 if i is located in the group she was allocated to at the beginning of the experiment, and 0 otherwise; $Group\ Contribution_{g[-i],t-1}$, the average contribution of the other members of i's current group in

[16]For ease of exposition, we estimate the models and present results for each of these outcomes separately. The findings reported below are generally similar if we fit panel dynamic simultaneous equations models, even though standard estimation approaches (e.g., Akashi and Kunitomo, 2012; Hsiao and Zhou, 2015) require ignoring the censoring of contributions and the dichotomous nature of migration and voting decisions.

[17]"Raw" parameter estimates – representing the effect of the predictors on the uncensored latent variable – are reported in the Appendix (Table A.1).

period $t-1$; and *abs(Group Size$_{g,t}$ - 5)*, the absolute value of the difference between 5 and the number of members in i's current group, measuring the impact of deviations from the "optimal" group size on individual contributions. The model also includes subject and period random effects to account for time-invariant individual heterogeneity and common temporal shocks affecting all subjects, along with session and group random intercepts to control for potential session effects (Fréchette, 2012) and contemporaneous correlation between same-group members (Poen, 2009).[18]

Individual contributions to the public account increase significantly when subjects belong to a group that implements sanctioning institutions. This result suggests participants generally understood the mechanism at play and responded to incentives, and reinforces the group-level evidence in support of Hypothesis 2.3. Also, everything else equal, average individual contributions are significantly higher when subjects play in their initial group. This finding again is consistent with incentives, because the returns to contributions to the public account are higher when participants remain in their initial group than when they are in the other group. Participants also tend to contribute more after observing higher contribution levels from their fellow group members in the previous period. On the other hand, i's contribution drops when the size of her group deviates from 5.

Column (2) repeats the specification in column (1), but accommodates possible asymmetric impacts that deviations from the optimal group size might have on subjects' contributions. We disaggregate *abs(Group Size$_{g,t}$ - 5)* into $\mathbb{1}$*(Group Size$_{g,t}$<5)(5 - Group Size$_{g,t}$)* and $\mathbb{1}$*(Group Size$_{g,t}$>5)(Group Size$_{g,t}$ - 5)*, where $\mathbb{1}(A)$ is the indicator function taking the value 1 if A is true, and 0 otherwise. The coefficients for both variables are negative and statistically significant, indicating that any deviation from 5 is systematically associated with a drop in individual contributions.

To account for autocorrelation and for subjects' tendency to adjust their contributions

[18] The results are generally similar for fixed-effects censored regression models estimated using Dhaene and Jochmans (2015)'s split-panel jackknife approach to correct for incidental parameter bias, as well as using Honoré (1992)'s semiparametric estimator. See Table A.2 in the Appendix.

Table 3
Determinants of individual contributions

	(1)	(2)	(3)	(4)
Is Punishing$_{g,t}$	13.86***	13.82***	12.93***	10.82***
	(0.29)	(0.29)	(0.29)	(1.30)
Initial Group$_{i,t}$	1.73***	1.77***	1.36***	0.51*
	(0.31)	(0.32)	(0.31)	(0.31)
Group Contribution$_{g[-i],t-1}$	0.13***	0.13***	0.03**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
abs(Group Size$_{g,t}$ *- 5)*	-0.36***		-0.29***	0.39***
	(0.10)		(0.09)	(0.10)
$\mathbb{1}$*(Group Size*$_{g,t}$*<5)(5 - Group Size*$_{g,t}$*)*		-0.47***		
		(0.16)		
$\mathbb{1}$*(Group Size*$_{g,t}$*>5)(Group Size*$_{g,t}$ *- 5)*		-0.33***		
		(0.11)		
Contribution$_{i,t-1}$			0.14***	0.12***
			(0.01)	(0.01)
Observations	7,250	7,250	7,250	7,250
Log likelihood	-24,993.46	-24,987.21	-23,777.47	-24,246.43

Notes. The table reports marginal effects for the covariates included in the panel models for individual contributions. The dependent variable is left-censored at 0 and right-censored at 50. Units of observation are individuals-per-period. All specifications include subject, period, group, and session random effects. Standard errors are presented in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

over time (Smith, 2013), column (3) incorporates the lag of the dependent variable among the regressors of the column (1) specification. The estimate for $Contribution_{i,t-1}$ indicates that roughly 14% of subject i's contribution in period t is explained by her contribution in the previous period. Nonetheless, the substantive conclusions emerging from the static models remain unchanged, although the association between $Contribution_{i,t}$ and the other explanatory variables – and in particular, the marginal effect of $Group\ Contribution_{g[-i],t-1}$ – becomes weaker.

Consistent with the results for *VM*, the estimates in column (4) – which replicates the column (3) specification for *NVM* – show again significant persistence in individuals' propensity to cooperate, which is further strengthened by sanctioning institutions. However, the previous contributions of other members of i's current group do not have a systematic influence on $Contribution_{i,t}$ when subjects are not allowed to choose the rule-set governing their interactions, whereas the impact of deviations from the optimal group size goes in the opposite direction.[19]

Migration. We now study the main determinants of subjects' migration decisions. Column (1) of Table 4 reports marginal effects obtained from a panel probit model fitted to data from our main treatment. The dependent variable is $Migration_{i,t}$, a dummy taking the value 1 if subject i moved between groups in period t, and 0 otherwise. We include as predictors of the moving decision the same set of explanatory variables used in the analysis of individual contributions, replacing $Group\ Contribution_{g[-i]}$ with $Contribution_{i,t}$ - $Group\ Contribution_{g[-i],t}$, the difference between i's contribution and the average contribution of the other group members in period t.[20] The model also incorporates subject-specific (correlated) random effects (Wooldridge, 2005, 2010) as well as period, group, and session random intercepts.[21]

All the covariates in the model have a statistically significant impact on $Migration_{i,t}$. All else equal, subjects are less likely to move if they are located in their initial group and if their current group has adopted a sanctioning institution. On the other hand, the larger the difference between i's contribution and the average contribution of other members of her group in period t, the higher the probability that she moves. Also, the probability that the average subject moves in any given period rises when the number of members of her group deviates from its optimal size.[22]

[19] To account for the potential endogeneity of covariates and allow for dynamic feedback from the outcome to the explanatory variables, Table A.3 in the Appendix replicates all the analyses in Table 3 using Fernández-Val and Vella (2011)'s two-step estimator for non-linear panel data models. The main results remain qualitatively similar, although deviations from the optimal group size are no longer significant in *VM*.

[20] Recall that individuals choose the amount of their contributions before deciding whether to move. This variable is not included in the models for individual contributions because of obvious collinearity problems in the dynamic specifications.

[21] Parameter estimates are reported in the Appendix (Table A.4). For robustness, we also fitted two-way (subject and period) fixed-effects probit models using the method developed by Fernández-Val and Weidner (2016). The estimates are aligned with those from the random effects models (see Table A.5 in the Appendix).

[22] Deviations both above and below 5 are statistically significant, although the marginal effect on migration is somewhat larger when the size of the group is below the optimal size.

Table 4
Determinants of individual migration decisions

	(1)	(2)	(3)	(4)	(5)
Is Punishing$_{g,t}$	-8.83***	-8.63***	-8.89***	-8.88***	-22.20***
	(1.15)	(1.09)	(1.15)	(1.15)	(2.74)
Initial Group$_{i,t}$	-23.20***	-23.15***	-23.21***	-21.94***	-18.01***
	(1.93)	(1.94)	(1.93)	(1.96)	(1.61)
Contribution$_{i,t}$ - *Group Contribution*$_{g[-i],t}$	3.18***	3.18***	3.15***	3.17***	1.86***
	(0.41)	(0.42)	(0.41)	(0.40)	(0.39)
abs(Group Size$_{g,t}$ *- 5)*	1.50***	1.49***	1.51***	1.51***	1.04***
	(0.30)	(0.28)	(0.27)	(0.29)	(0.36)
Vote Different from the Group$_{i,t-1}$		6.48***			
		(1.94)			
Vote Different from the Group$_{i,t-1}$ *under Rule-Set 1*			4.50	4.31	
			(3.33)	(3.32)	
Vote Different from the Group$_{i,t-1}$ *under Rule-Set 2*			7.83***	8.08***	
			(2.47)	(2.51)	
Pre-Voting Period$_t$ - *Initial Group*$_{i,t}$				-0.79	
				(0.92)	
Pre-Voting Period$_t$ - *Outside Initial Group*$_{i,t}$				6.89**	
				(3.45)	
Observations	7,250	7,250	7,250	7,250	7,250
Log likelihood	-2,091.33	-2,082.14	-2,080.44	-2,078.04	-2,568.16

Notes. The table reports the change in the probability of moving (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%.

Column (2) incorporates a dummy variable, *Vote Different from the Group*$_{i,t-1}$, measuring whether i's institutional choice in the most recent voting period differed from the decision of the majority of her group. We observe that the probability of migrating increases when the group did not adopt subjects' preferred rule-set in the most recent voting round. To determine whether this positive effect is contingent on the institutional setting implemented, column (3) adds an interaction between *Vote Different from the Group*$_{i,t-1}$ and *Is Punishing*$_{g,t}$ to the previous specification. We can therefore compute the marginal effects of *Vote Different from the Group*$_{i,t-1}$ *under Rule-Set 1* (i.e., when *Is Punishing*$_{g,t}$=0) and *Vote Different from the Group*$_{i,t-1}$ *under Rule-Set 2* (when *Is Punishing*$_{g,t}$=1). Although the probability of moving is always higher on average for "dissenters" than for subjects who agreed with the majority decision, only those who voted against sanctioning in groups that adopted centralized punishment are significantly more likely to migrate.

In column (4), we add an interaction between *Pre-Voting Period*$_t$ – a dummy for periods immediately before each voting round – and *Initial Group*$_{i,t}$, which allows us to assess whether subjects in/out of their initial group are more or less likely to move right before groups choose their institutions. Specifically, we calculate the marginal effects of *Pre-Voting Period*$_t$ when subjects are currently located in their initial group (setting *Initial Group*$_{i,t}$=1 for all i) and when subjects are located outside their initial group (i.e., when *Initial Group*$_{i,t}$=0 for all i). In accordance with Hypothesis 2.5, we find that subjects located outside their initial group are more likely to move back right before a voting round – presumably to affect the outcome of the vote – than in other periods. By contrast, for individuals located in their initial group, no significant differences exist in the probability of moving in pre-voting vis-à-vis other periods. Also, we find no evidence of individuals being more likely to return to their original group right after a voting round, once the institutional choice is known.[23]

For comparison, column (5) repeats the specification in column (1) using data from *NVM*. The marginal effects of the covariates are qualitatively similar to those for *VM*, although the

[23]Additionally, we estimated models including the lagged dependent variable among the regressors, but we find no evidence of state dependence in migration decisions.

impact of the sanctioning institution on the probability of moving is one and a half times larger – in absolute value – when subjects are not allowed to vote on the rule-set.[24]

Voting. Lastly, we examine the determinants of subjects' voting decisions. Table 5 reports marginal effects computed from random effects panel probit models fitted to data from the five voting rounds included in *VM*.[25] The dependent variable, $Vote_{i,t}$, equals 1 (0) if i voted for (against) the sanctioning institution in period t.

The explanatory variables in column (1) are essentially the same as those of previous individual-level analyses, with *Is Punishing* now lagged one period to capture whether Rule-Set 2 is in place in i's group before the current voting round. As stated in Remark 4, the probability of voting in favor of the sanctioning institution is systematically higher for subjects who remain in the group to which they were originally allocated. The average propensity to support sanctions also rises for subjects located in groups that already implement sanctions before t, and when the number of members in i's current group differs from 5.[26] These last two effects are, however, only marginally significant.

Column (2) adds a linear time trend to the column (1) specification and shows the probability of voting in favor of sanctions grows by 4.8 percentage points on average between (consecutive) voting rounds, after controlling for other observed and unobserved factors.[27] This finding reflects the patterns observed in the data and provides further support for Hypothesis 2.2, with the proportion of subjects voting for sanctions going from 41% in the first voting period to 62% in the last round.

The steady increase in the probability of voting for institutions over the course of the experiment is indicative of strong persistence in individual voting decisions. This result is

[24] Table A.6 in the Appendix replicates the analyses in Table 4 using Fernández-Val and Vella (2011)'s two-step estimator to account for potential endogeneity of the explanatory variables and for dynamic feedback from $Migration_{i,t}$ to the predetermined regressors. Although the marginal effect of *abs(Group Size*$_{g,t}$ *- 5)* is never significant, the estimates for the other covariates remain similar.

[25] Parameter estimates are reported in Table A.7 of the Appendix. For completeness, Table A.8 presents estimates from bias-corrected fixed-effects panel probit models (Fernández-Val and Weidner, 2016).

[26] In particular, $\mathbb{1}$*(Group Size>5)(Group Size-5)* - but not of $\mathbb{1}$*(Group Size<5)(5-Group Size)* - has a positive and significant marginal effect on $Pr(Vote_{i,t} = 1)$.

[27] The significant increase in $Pr(Vote_{i,t} = 1)$ across periods is also observed using a quadratic time trend.

Table 5
Determinants of individual voting decisions.

	(1)	(2)	(3)
Is Punishing$_{g,t-1}$	5.39*	4.77	-3.10
	(3.09)	(3.09)	(3.61)
Initial Group$_{i,t}$	9.69***	9.41***	7.61**
	(3.57)	(3.55)	(3.76)
Contribution$_{i,t}$ - *Group Contribution*$_{g[-i],t}$	-1.65	-1.64	-2.67
	(1.41)	(1.41)	(1.64)
abs(Group Size$_{g,t}$ *- 5)*	1.52*	1.48**	1.81**
	(0.79)	(0.73)	(0.91)
Time Trend		4.80***	3.14***
		(0.76)	(1.19)
Vote$_{i,t-1}$			29.79***
			(5.39)
Observations	1,250	1,250	1,000
Log likelihood	-612.38	-605.16	-457.79

Notes. The table reports the change in the probability of voting for sanctions (in percentage points) associated with a change in the covariates. Units of observation are individuals-per-voting period. All specifications include subject, period, group, and session random effects; columns (2)-(3) also include a linear time trend. Standard errors are presented in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

confirmed in Column (3), which adds the lagged dependent variable to the set of covariates. The probability of voting for sanctions rises by almost 30% for a subject who already voted in favor of Rule-Set 2 in the previous voting round. The marginal effect of *Is Punishing*$_{g,t-1}$ ceases to be significant once we control for state dependence and for the growing propensity to support sanctions over time.[28]

6 Conclusions

This paper analyzes the effect of the endogenous choice of sanctioning institutions on individuals' levels of cooperation and migration decisions. In our experiment, subjects are

[28] Table A.9 in the Appendix replicates the analyses in Table 5 accounting for covariate endogeneity through a control-function approach for random effects panel probit models (Giles and Murtazashvili, 2013). Unlike Fernández-Val and Vella (2011)'s method, this approach does not rely on large-T asymptotics, and is thus better suited for inference with only five voting periods. The estimates are in line with those in Table 5, although the marginal effect for *Is Punishing*$_{g,t-1}$ remains positive and significant in all specifications. We also applied special regressor estimators (Dong and Lewbel, 2015) and obtained similar results.

assigned to one of two groups and face a repeated social dilemma. Individuals can move between groups and, in the main treatment, they vote at fixed intervals on whether to implement a formal (centralized) sanctioning institution in the group in which they are located.

To study the effect of location choices on both individuals and societies, we introduce two features to our setting. First, subjects have group-specific preferences in that, all else equal, they enjoy higher payoffs if located in the group to which they are initially assigned. This feature accounts for the individual cost of migrating, which is generated by family and social ties. Second, to capture the cost imposed on societies by sizable migration outflows, we make the return from the public good (weakly) increasing in the group size. These two features imply, all else equal, an efficient partition of individuals across groups: at optimum, each individual contributes to the public good in the group to which she is initially assigned.

We compare this treatment with one in which only one group is exogenously endowed with sanctioning institutions. When subjects cannot vote over the sanctioning regime, because sanctions positively affect contributions to the public good, the group endowed with sanctioning institutions (resp., the sanction-free group) tends to become overpopulated (resp., depopulated). Even though sanctions ensure high overall contributions to the public good, this migration flow causes payoff inequality both between groups and within the overpopulated group. By contrast, when subjects can vote over the sanctioning regime, groups increasingly implement sanctions over time, provide (more) efficient levels of the public good, and coordinate on the efficient partition of individuals across groups. Subjects' efficient location choices significantly increase overall payoffs and reduce payoff inequality between and within groups. At the individual level, subjects migrate less when voting over institutions is allowed; when migrating, they tend to move back to their initial group in order to cast their ballot and influence their initial group's institutional choice.

We conclude with two remarks. First, in our laboratory experiment, subjects manage remarkably well to self-organize in different dimensions. Even though complex individual decisions – such as where to locate – can only be imperfectly approximated in a laboratory setting, this controlled environment allows us to take a close look at the composite relationship between institutions, cooperation, and migration. This analysis, in turn, improves our understanding of groups' ability to select efficient outcomes. Second, our results seem to suggest formal institutions are sufficiently strong to stabilize migration patterns. In reality, however, informal sanctioning institutions – and their interaction with formal ones – also play a fundamental role in societies' ability to solve social dilemmas. Whether peer-to-peer sanctioning institutions also adapt to and stabilize migration patterns to achieve efficient outcomes remains an open question, which we leave for future research.

References

Acemoglu, Daron, Simon Johnson and James Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *The American Economic Review* 91(5):1369–1401.

Acemoglu, Daron, Simon Johnson and James Robinson. 2005. "The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth." *The American Economic Review* 95(3):546–579.

Akashi, Kentaro and Naoto Kunitomo. 2012. "Some Properties of the LIML Estimator in a Dynamic Panel Structural Equation." *Journal of Econometrics* 166(2):167–183.

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

Angrist, Joshua D. 2001. "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice." *Journal of Business & Economic Statistics* 19(1):2–28.

Baldassarri, Delia and Guy Grossman. 2011. "Centralized Sanctioning and Legitimate Authority Promote Cooperation in Humans." *Proceedings of the National Academy of Sciences* 108(27):11023–11027.

Beine, Michel, Frédéric Docquier and Hillel Rapoport. 2007. "Measuring International Skilled Migration: a New Database Controlling for Age of Entry." *The World Bank Economic Review* 21(2):249–254.

Charness, Gary and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* pp. 817–869.

Chaudhuri, Ananish. 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments: a Selective Survey of the Literature." *Experimental Economics* 14(1):47–83.

Conley, John P. and Hideo Konishi. 2002. “Migration-proof Tiebout equilibrium: Existence and Asymptotic Efficiency.” *Journal of Public Economics* 86(2):243–262.

De Haas, Hein. 2010. “Migration and Development: a Theoretical Perspective.” *International Migration Review* 44(1):227–264.

Denant-Boemont, Laurent, David Masclet and Charles N. Noussair. 2007. “Punishment, Counterpunishment and Sanction Enforcement in a Social Dilemma Experiment.” *Economic theory* 33(1):145–167.

Dhaene, Geert and Koen Jochmans. 2015. “Split-panel Jackknife Estimation of Fixed-effect Models.” *Review of Economic Studies* 82(3):991–1030.

Dinkelman, Taryn and Sam Schulhofer-Wohl. 2015. “Migration, Congestion Externalities, and the Evaluation of Spatial Investments.” *Journal of Development Economics* 114:189–202.

Dong, Yingying and Arthur Lewbel. 2015. “A Simple Estimator for Binary Choice Models with Endogenous Regressors.” *Econometric Reviews* 34(1-2):82–105.

Dreber, Anna, David G. Rand, Drew Fudenberg and Martin A. Nowak. 2008. “Winners Don’t Punish.” *Nature* 452(7185):348–351.

Ertan, Arhan, Talbot Page and Louis Putterman. 2009. “Who to punish? Individual Decisions and Majority Rule in Mitigating the Free Rider Problem.” *European Economic Review* 53(5):495–511.

Fehr, Ernst and Klaus M. Schmidt. 1999. “A Theory of Fairness, Competition, and Cooperation.” *Quarterly Journal of Economics* 114(3):817–868.

Fehr, Ernst and Simon Gächter. 2000. “Cooperation and Punishment in Public Goods Experiments.” *The American Economic Review* 90(4):980–994.

Fehr, Ernst and Simon Gächter. 2002. “Altruistic Punishment in Humans.” *Nature* 415(6868):137–140.

Fernández-Val, Iván and Francis Vella. 2011. “Bias Corrections for Two-step Fixed Effects Panel Data Estimators.” *Journal of Econometrics* 163(2):144–162.

Fernández-Val, Iván and Martin Weidner. 2016. “Individual and Time Effects in Nonlinear Panel Models with Large N, T.” *Journal of Econometrics* 192(1):291–312.

Fréchette, Guillaume R. 2012. “Session-effects in the Laboratory.” *Experimental Economics* 15(3):485–498.

Friedberg, Rachel M. 2001. “The Impact of Mass Migration on the Israeli Labor Market.” *Quarterly Journal of Economics* 116(4):1373–1408.

Gächter, Simon and Benedikt Herrmann. 2009. “Reciprocity, Culture and Human Cooperation: Previous Insights and a New Cross-cultural Experiment.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1518):791–806.

Giles, John and Irina Murtazashvili. 2013. “A Control Function Approach to Estimating Dynamic Probit Models with Endogenous Regressors.” *Journal of Econometric Methods* 2(1):69–87.

Greenberg, Joseph and Shlomo Weber. 1986. “Strong Tiebout Equilibrium under Restricted Preferences Domain.” *Journal of Economic Theory* 38(1):101–117.

Greene, William H. 2012. *Econometric Analysis*. Essex, UK: Pearson.

Gross, Jörg, Zsombor Z. Méder, Sanae Okamoto-Barth and Arno Riedl. 2016. “Building the Leviathan – Voluntary Centralisation of Punishment Power Sustains Cooperation in Humans.” *Scientific Reports* 6.

Guala, Francesco. 2012. “Reciprocity: Weak or Strong? What Punishment Experiments Do (and Do Not) Demonstrate.” *Behavioral and Brain Sciences* 35(01):1–15.

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

Gürerk, Özgür, Bernd Irlenbusch and Bettina Rockenbach. 2009. "Motivating teammates: The Leader's Choice between Positive and Negative Incentives." *Journal of Economic Psychology* 30(4):591–607.

Gürerk, Özgür, Bernd Irlenbusch and Bettina Rockenbach. 2014. "On Cooperation in Open Communities." *Journal of Public Economics* 120:220–230.

Herrmann, Benedikt, Christian Thöni and Simon Gächter. 2008. "Antisocial Punishment Across Societies." *Science* 319(5868):1362–1367.

Hill, John K. 1987. "Immigrant Decisions Concerning Duration of Stay and Migratory Frequency." *Journal of Development Economics* 25(1):221–234.

Honoré, Bo E. 1992. "Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects." *Econometrica* 60(3):533–565.

Hsiao, Chen and Qiankun Zhou. 2015. "Statistical Inference for Panel Dynamic Simultaneous Equations models." *Journal of Econometrics* 189(2):383–396.

Jones, Eric. 1981. *The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia*. Cambridge, UK: Cambridge University Press.

Katseli, Louka T., Robert E. B. Lucas and Theodora Xenogiani. 2006. "Effects of Migration on Sending Countries: What do We Know?" OECD Development Centre Working Paper No. 250.

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

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

Niedomysl, Thomas and Jan Amcoff. 2011. "Why Return Migrants Return: Survey Evidence on Motives for Internal Return Migration in Sweden." *Population, Space and Place* 17(5):656–673.

Nikiforakis, Nikos. 2008. "Punishment and Counter-punishment in Public Good Games: Can We really Govern Ourselves?" *Journal of Public Economics* 92(1):91–112.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance.* Cambridge, UK: Cambridge University Press.

North, Douglass C. and Robert Paul Thomas. 1973. *The Rise of the Western World: A New Economic History.* Cambridge University Press.

North, Douglass Cecil. 1981. *Structure and Change in Economic History.* Cambridge, UK: W. W. Norton & Co.

Ostrom, Elinor, James Walker and Roy Gardner. 1992. "Covenants with and without a Sword: Self-Governance Is Possible." *American Political Science Review* 86(2):404–417.

Pesarin, Fortunato and Luigi Salmaso. 2010. *Permutation Tests for Complex Data: Theory, Applications and Software.* Chichester, UK: Wiley & Sons.

Poen, Eva. 2009. "The Tobit Model with Feedback and Random Effects: A Monte-Carlo Study." CeDEx Discussion Paper Series 2009-14, University of Nottingham.

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

Sigmund, Karl, Hannelore De Silva, Arne Traulsen and Christoph Hauert. 2010. "Social Learning Promotes Institutions for Governing the Commons." *Nature* 466(7308):861–863.

Smith, Alexander. 2013. "Estimating the Causal Effect of Beliefs on Contributions in repeated public good games." *Experimental Economics* 16(3):414–425.

Stock, James H. and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regressions". In *Identification and Inference for Econometric Models*, ed. Donald W. Andrews and James H. Stock. New York, NY: Cambridge University Press, pp. 80–108.

Sutter, Matthias, Stefan Haigner and Martin G. Kocher. 2010. "Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations." *The Review of Economic Studies* 77(4):1540–1566.

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

Traulsen, Arne, Torsten Röhl and Manfred Milinski. 2012. "An Economic Experiment Reveals that Humans Prefer Pool Punishment to Maintain the Commons." *Proceedings of the Royal Society of London B: Biological Sciences* 279(1743):3716–3721.

Van der Heijden, Eline, Jan Potters and Martin Sefton. 2009. "Hierarchy and Opportunism in Teams." *Journal of Economic Behavior & Organization* 69(1):39–50.

Wooldridge, Jeffrey M. 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, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Yamagishi, Toshio. 1986. "The Provision of a Sanctioning System as a Public Good." *Journal of Personality and Social Psychology* 51(1):110.

Zhang, Boyu, Cong Li, Hannelore De Silva, Peter Bednarik and Karl Sigmund. 2014. "The Evolution of Sanctioning Institutions: an Experimental Approach to the Social Contract." *Experimental Economics* 17(2):285–303.

Appendix

A. Experimental Instructions

A.1. First Stage

You will be in a group of 5 people. You will receive 50 tokens (converted to cash at the end of the experiment) and will choose to allocate some portion of these to your own individual account and some portion to the group account. Your final payoffs will be as follows:

50 – (tokens you allocate to group account) + (1/5) × 2.250× (sum of tokens allocated by all in group to group account)

At the end of the experiment tokens will be converted into GBP at the rate of 30 tokens = 1 GBP.

A.2. Second Stage: *Voting-Moving* Treatment

There will be a total of 30 periods in the second activity. We will explain carefully what you have to do in this second part.

Allocations: In each period, you will be in a group of some size (you could be also by yourself). In each period, you will receive 50 tokens (converted to cash at the end of the experiment) and will choose to allocate some portion of these to your own individual account and some portion to the group account. Your final payoffs will be as follows:

50 – (tokens you allocate to group account) + $(1/n)\times$ Factor × (sum of tokens allocated by all in group to group account)

where n is the number of members of the group. The Factor that multiplies the sum of the tokens is explained below.

Groups: In the first period you will be assigned to a five people group (including yourself). You can be assigned to either GROUP A or GROUP B. At the end of each period you will have the chance of deciding whether you want to leave your group. We will explain in more detail how you do this.

Group size and multiplying Factor: The total number of tokens that end up in the group account is multiplied by a Factor. The table below shows the values of the Factor depending on: *(i)* the group you are in and *(ii)* the group size.

Group Size (n)	**Factor if in the Initial Group**	**Factor if not in the Initial Group**
1	1.00	1.00
2	1.50	1.15
3	1.75	1.40
4	2.00	1.65
5	2.25	1.90
6	2.25	1.90
7	2.25	1.90
8	2.25	1.90
9	2.25	1.90
10	2.25	1.90

Each token allocated to the group account yields a return higher than one unit and all members of the group share this return equally.

Note that the multiplying factor increases as the group size increases. The increment of the factor with the group size will be capped at 5 members in the group. If the size of the group goes above that, the factor will remain the same as if the size of the group was 5.

There will be two different factors depending on whether you are in your **initial** group or if you decided to switch groups. Given the size of the group, the factor will be always larger if you are in your **initial** group than if you are in the other group. Thus, for the same amount of tokens allocated to the group account, and for the same size of the group, tokens allocated to the group account will generate more money if you are in your **initial** group.

Each token allocated to one's private account pays one unit to that individual, independently of the group size.

EXAMPLE: Consider the case in which one participant belongs to **Group A** and her initial group was **Group A**. If in a five-person group the total number of tokens in the group account is 100, that participant receives a payoff of $(100 \times 2.250)/5{=}45$ tokens from the group account (plus the number of tokens allocated to her private account). By comparison, suppose the participant is in **Group B** and her initial group was **Group A**. If in a five-person group the total number of tokens in the group account is 100, then, that participant receives a payoff of $(100 \times 1.900)/5{=}38$ tokens from the group account (plus the number of tokens allocated to her private account).

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

RULE SET 1 (no points reduction): In Rule Set 1, earnings are determined in exactly the same way as explained in the **Allocations** section. So, payoffs will be computed as follows

50 – (points you allocate to group account) $+$ $(1/n)\times$ Factor $\times$ (sum of points allocated by all in group to group account)

RULE SET 2 (automatic points reduction): In Rule Set 2, each individual pays a fixed fee of 15 tokens 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 tokens allocated to the private account in that period. Payoffs in each period are calculated as follows:

50 – (points you allocate to group account) + $(1/n)\times$ Factor $\times$ (sum of points allocated by all in group to group account) – 0.8 $\times$(points you allocate to private account) – 15

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 tokens to the group account and keeps 30 tokens for the private account, given RULE SET 2, that participant would have to pay a fine of $0.8 \times 30=24$ tokens and then, she would keep 6 tokens out of 30.

Payment rules in first 5 periods: If you were assigned to Group A, Rule Set 1 will apply for the first 5 periods. If you were assigned to group B, Rule set 2 will apply for the first 5 periods.

Group determination: At the end of each period, you will observe: *(i)* the average contribution of your current group, *(ii)* the average contribution of the other group, *(iii)* the rule set of your current group, and *(iv)* the rule set of the other group. Then, you will decide whether you wish to leave the group and move to the other one.

Voting for rules: Initial rules for Groups A and B will apply only for the first 5 periods. Every five periods you will have the chance to vote on whether you want Rule Set 1 or Rule Set 2 to be applied to the group you currently belong to. Note that when you vote, you will vote to establish a rule in the group that you belong to in the voting period.

Payment: At the end of the experiment tokens will be converted into GBP at the rate of 30 tokens = 1GBP. You will be paid only for 3 randomly chosen periods.

SUMMARY

1. There will be a total of 30 periods.
2. You will begin by being in a five-people group.
3. In each period, you will allocate 50 tokens between a private and a group account.
4. The returns from the group account depend on the size of the group and on whether you are participating in your initial group or not.
5. You can decide at the end of the period whether you want to switch groups or not.
6. There are two payment rules. Every five period you will vote which of the two rules you prefer.
7. You will be paid only for 3 rounds (randomly chosen).

A.3. Second Stage: *No-Voting-Moving* Treatment

There will be a total of 30 periods in the second activity. We will explain carefully what you have to do in this second part.

Allocations: In each period, you will be in a group of some size (you could be also by yourself). In each period, you will receive 50 tokens (converted to cash at the end of the experiment) and will choose to allocate some portion of these to your own individual account and some portion to the group account. Your final payoffs will be as follows:

50 – (tokens you allocate to group account) + $(1/n)\times$ Factor $\times$ (sum of tokens allocated by all in group to group account)

where n is the number of members of the group. The Factor that multiplies the sum of the tokens is explained below.

Groups: In the first period you will be assigned to a five people group (including yourself). You can be assigned to either GROUP A or GROUP B. At the end of each period you will have the chance of deciding whether you want to leave your group. We will explain in more detail how you do this.

Group size and multiplying Factor: The total number of tokens that end up in the group account is multiplied by a Factor. The table below shows the values of the Factor depending on: *(i)* the group you are in and *(ii)* the group size.

Group Size (n)	**Factor if in the Initial Group**	**Factor if not in the Initial Group**
1	1.00	1.00
2	1.50	1.15
3	1.75	1.40
4	2.00	1.65
5	2.25	1.90
6	2.25	1.90
7	2.25	1.90
8	2.25	1.90
9	2.25	1.90
10	2.25	1.90

Each token allocated to the group account yields a return higher than one unit and all members of the group share this return equally.

Note that the multiplying factor increases as the group size increases. The increment of the factor with the group size will be capped at 5 members in the group. If the size of the group goes above that, the factor will remain the same as if the size of the group was 5.

There will be two different factors depending on whether you are in your **initial** group or if you decided to switch groups. Given the size of the group, the factor will be always larger if you are in your **initial** group than if you are in the other group. Thus, for the same amount of tokens allocated to the group account, and for the same size of the group, tokens allocated to the group account will generate more money if you are in your **initial** group.

Each token allocated to one's private account pays one unit to that individual, independently of the group size.

EXAMPLE: Consider the case in which one participant belongs to **Group A** and her initial group was **Group A**. If in a five-person group the total number of tokens in the group account is 100, that participant receives a payoff of $(100 \times 2.250)/5=45$ tokens from the group account (plus the number of tokens allocated to her private account). By comparison, suppose the participant is in **Group B** and her initial group was **Group A**. If in a five-person group the total number of tokens in the group account is 100, then, that participant receives a payoff of $(100 \times 1.900)/5=38$ tokens from the group account (plus the number of tokens allocated to her private account).

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

RULE SET 1 (no points reduction): In Rule Set 1, earnings are determined in exactly the same way as explained in the **Allocations** section. So, payoffs will be computed as follows

50 – (points you allocate to group account) $+ (1/n)\times$ Factor $\times$ (sum of points allocated by all in group to group account)

RULE SET 2 (automatic points reduction): In Rule Set 2, each individual pays a fixed fee of 15 tokens 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 tokens allocated to the private account in that period. Payoffs in each period are calculated as follows:

$50 - $ (points you allocate to group account) $+ (1/n)\times$ Factor $\times$ (sum of points allocated by all in group to group account) $- 0.8 \times$(points you allocate to private account) $- 15$

<u>EXAMPLE:</u> 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 tokens to the group account and keeps 30 tokens for the private account, given RULE SET 2, that participant would have to pay a fine of $0.8 \times 30=24$ tokens and then, she would keep 6 tokens out of 30.

Payment rules: If you were assigned to Group A, Rule Set 1 will apply. If you were assigned to group B, Rule set 2 will apply.

Group determination: At the end of each period, you will observe: *(i)* the average contribution of your current group, *(ii)* the average contribution of the other group, *(iii)* the rule set of your current group, and *(iv)* the rule set of the other group. Then, you will decide whether you wish to leave the group and move to the other one.

Payment: At the end of the experiment tokens will be converted into GBP at the rate of 30 tokens = 1GBP. You will be paid only for 3 randomly chosen periods.

SUMMARY

1. There will be a total of 30 periods.
2. You will begin by being in a five-people group.
3. In each period, you will allocate 50 tokens between a private and a group account.
4. The returns from the group account depend on the size of the group and on whether you are participating in your initial group or not.

5. You can decide at the end of the period whether you want to switch groups or not.

6. There are two payment rules.

7. You will be paid only for 3 rounds (randomly chosen).

B. Additional Results

Table A.1
Parameter Estimates - Censored Regression Models for Contributions

	(1)	(2)	(3)	(4)
Intercept	18.84***	18.93***	15.15***	18.94***
	(1.10)	(1.13)	(1.10)	(1.82)
Is Punishing$_{g,t}$	33.19***	33.10***	31.13***	26.24***
	(0.77)	(0.76)	(0.74)	(3.12)
Initial Group$_{i,t}$	4.14***	4.25***	3.28***	1.23*
	(0.75)	(0.78)	(0.74)	(0.74)
Group Contribution$_{g[-i],t-1}$	0.30***	0.30***	0.08***	-0.03
	(0.02)	(0.02)	(0.03)	(0.03)
abs(Group Size$_{g,t}$ *- 5)*	-0.86***		-0.70***	0.95***
	(0.24)		(0.22)	(0.25)
$\mathbb{1}$*(Group Size*$_{g,t}$*<5)(5 - Group Size*$_{g,t}$*)*		-1.13***		
		(0.39)		
$\mathbb{1}$*(Group Size*$_{g,t}$*>5)(Group Size*$_{g,t}$ *- 5)*		-0.78***		
		(0.25)		
Contribution$_{i,t-1}$			0.35***	0.28***
			(0.02)	(0.02)
Observations	7,250	7,250	7,250	
Log likelihood	-24,993.46	-24,987.21	-23,777.47	-24,246.43

Notes. The table reports "raw" parameter estimates for the static (columns 1-2) and dynamic (column 3-4) panel data models used to compute the marginal effects reported in Table 3. Columns (1)-(3) are fitted to data from the *Voting-Moving* treatment, whereas column (4) uses data from the *No-Voting-Moving* condition. 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 A.2

Marginal Effects of Covariates on Individual Contributions
Estimated from Fixed-Effects Censored Regression Models

	(1)	(2)	(3)	(4)	(5)	(6)
Is Punishing$_{g,t}$	14.77***	14.74***	13.42***	14.23***	19.46***	19.29***
	(0.30)	(0.30)	(0.29)	(0.33)	(1.10)	(1.09)
Initial Group$_{i,t}$	1.41**	1.53**	1.24***	0.19	1.77**	2.13**
	(0.33)	(0.35)	(0.32)	(0.33)	(0.81)	(0.82)
Group Contribution$_{g[-i],t-1}$	0.13***	0.13***	0.03**	-0.02*	0.15***	0.14***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
abs(Group Size$_{g,t}$ *- 5)*	-0.46***		-0.42***	0.37***	-0.42***	
	(0.10)		(0.10)	(0.10)	(0.14)	
$\mathbb{1}$*(Group Size*$_{g,t}$$<$*5)(5 - Group Size*$_{g,t}$*)*		-0.58***				-1.07***
		(0.18)				(0.34)
$\mathbb{1}$*(Group Size*$_{g,t}$$>$*5)(Group Size*$_{g,t}$ *- 5)*		-0.44**				-0.31**
		(0.10)				(0.15)
Contribution$_{i,t-1}$			0.16***	0.14***		
			(0.01)	(0.01)		
Observations	7,250	7,250	7,250	7,250	7,250	7,250

Notes. The table reports marginal effects of the covariates on subjects' contributions, obtained from fixed-effects panel data censored regression models. Columns (1)-(4) replicate the analyses in columns (1)-(4) of Table 3, using the split-panel jackknife estimation approach proposed by Dhaene and Jochmans (2015) to correct for incidental parameter bias. For robustness, Columns (5) and (6) also fit the first two static models using Honoré (1992)'s semiparametric estimator. Columns (1)-(3) and (5)-(6) are fitted to data from the *Voting-Moving* treatment, while column (4) uses data from the *No-Voting-Moving* condition. All specifications include subject fixed-effects. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table A.3
Determinants of Individual Contributions - Accounting for Covariate Endogeneity Using Fernández-Val and Vella (2011)'s Two-Step Estimator

	(1)	(2)	(3)	(4)
Is Punishing$_{g,t}$	12.91***	12.89***	10.92***	14.43***
	(1.41)	(1.67)	(2.47)	(0.57)
Initial Group$_{i,t}$	2.72**	2.76**	2.33*	4.57***
	(1.25)	(1.30)	(1.35)	(0.38)
Group Contribution$_{g[-i],t-1}$	0.16***	0.16***	0.06***	-0.01
	(0.02)	(0.03)	(0.02)	(0.01)
abs(Group Size$_{g,t}$ *- 5)*	-0.46		-0.48	0.31***
	(0.39)		(0.39)	(0.12)
$\mathbb{1}$*(Group Size*$_{g,t}$*<5)(5 - Group Size*$_{g,t}$*)*		-0.45		
		(0.40)		
$\mathbb{1}$*(Group Size*$_{g,t}$*>5)(Group Size*$_{g,t}$ *- 5)*		-0.44		
		(0.46)		
Contribution$_{i,t-1}$			0.17***	0.16***
			(0.03)	(0.01)
Observations	7,250	7,250	7,250	7,250

Notes. The table replicates the analyses in Table 3, reporting marginal effects of the covariates on subjects' contributions estimated using Fernández-Val and Vella (2011)'s two-step (control function) approach for non-linear panel data models. Columns (1)-(3) are fitted to data from the *Voting-Moving* treatment, while column (4) uses data from the *No-Voting-Moving* condition. These specifications treat *Is Punishing*$_{g,t}$ (in the *Voting-Moving* treatment), *Initial Group*$_{i,t}$ and the measures of deviations from the optimal group size as endogenous, and allow for dynamic feedback from the dependent variable to the predictors. The control variables are the (generalized) residuals from the reduced-form models for the endogenous variables, specified as functions of pre-determined regressors - their own lagged values, the lagged average contribution of other members of i's group, the lagged values of $R_h(n)$, indicators capturing whether subjects' decision in the most recent voting round coincided with their/the other group's institutional choice (in columns 1-3), dummies for periods right before and after a voting round (in columns 1-3), and the lag of *Is Punishing*$_{g,t}$ (in column 4) - along with subject fixed effects. All specifications in the table include subject fixed-effects; although time dummies (or time trends) are not covered by Fernández-Val and Vella (2011)' regularity conditions, adding them does not alter the results. Standard errors are reported in parentheses; we combined split-panel jackknife estimation and (panel) bootstrapping to correct for incidental parameter bias and adjust for generated regressors (Dhaene and Jochmans, 2015). Significance levels: *** at 1%, ** at 5%, * at 10%. The first-step (reduced-form) estimates are available from the authors upon request.

Table A.4

Parameter Estimates - Probit Panel Models for Migration

	(1)	(2)	(3)	(4)	(5)
Intercept	-0.91***	-0.95***	-0.99***	-1.02***	-0.25
	(0.18)	(0.19)	(0.19)	(0.19)	(0.19)
Is $Punishing_{g,t}$	-0.56***	-0.55***	-0.58***	-0.58***	-1.00***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.11)
Initial $Group_{i,t}$	-1.10***	-1.11***	-1.10***	-1.06***	-0.82***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)
$Contribution_{i,t}$ - *Group* $Contribution_{g[-i],t}$	0.21***	0.21***	0.21***	0.21***	0.10***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)
abs(Group $Size_{g,t}$ *- 5)*	0.12***	0.13***	0.13***	0.13***	0.06***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Vote Different from the $Group_{i,t-1}$		0.37***	0.20	0.20	
		(0.10)	(0.14)	(0.14)	
Vote Different from the $Group_{i,t-1} \times$ *Is* $Punishing_{g,t}$			0.32	0.33*	
			(0.19)	(0.19)	
Pre-Voting $Period_t$				0.22**	
				(0.11)	
Pre-Voting $Period_t \times$ *Initial* $Group_{i,t}$				-0.29**	
				(0.14)	
Observations	7,250	7,250	7,250	7,250	7,250
Log likelihood	-2,091.33	-2,082.14	-2,080.44	-2,078.04	-2,568.16

Notes. The table reports "raw" parameter estimates for the panel probit models used to compute the marginal effects reported in Table 4. Columns (1)-(4) are fitted to data from the *Voting-Moving* treatment, while column (5) uses data from the *No-Voting-Moving* condition. 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 A.5
Parameter Estimates - Fixed-Effects Probit Panel Models for Migration

	(1)	(2)	(3)	(4)	(5)
Is Punishing$_{g,t}$	-0.65***	-0.65***	-0.67***	-0.64***	-0.85***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Initial Group$_{i,t}$	-0.93***	-0.93***	-0.93***	-0.89***	-0.68***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)
Contribution$_{i,t}$ - *Group Contribution*$_{g[-i],t}$	0.20***	0.20***	0.20***	0.20***	0.10***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
abs(Group Size$_{g,t}$ *- 5)*	0.11***	0.11***	0.11***	0.09***	0.06***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Vote Different from the Group$_{i,t-1}$		0.37***	0.18	0.17	
		(0.12)	(0.16)	(0.14)	
Vote Different from the Group$_{i,t-1}$ $\times$ *Is Punishing*$_{g,t}$			0.37*	0.34*	
			(0.20)	(0.20)	
Pre-Voting Period$_t$				0.20*	
				(0.11)	
Pre-Voting Period$_t$ $\times$ *Initial Group*$_{i,t}$				-0.28**	
				(0.14)	
Observations	5,336	5,336	5,336	5,336	5,626
Log likelihood	-1,850.55	-1,845.60	-1,843.61	-1,856.93	-2,307.12

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. Columns (1)-(4) are fitted to data from the *Voting-Moving* treatment, while column (5) uses data from the *No-Voting-Moving* condition. Difference in number of observations vis-à-vis Tables 4 and A.4 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 fixed-effects. Columns (1)-(3) and (5) also include period fixed-effects, which must be dropped in column (4) in order to recover the coefficient for *Pre-Voting Period*$_t$. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table A.6

Determinants of Individual Migration Decisions - Accounting for Covariate Endogeneity using Fernández-Val and Vella (2011)'s Two-Step Estimator

	(1)	(2)	(3)	(4)	(5)
Is Punishing$_{g,t}$	-10.07***	-10.02***	-9.94***	-9.86***	-17.37***
	(2.50)	(2.52)	(2.51)	(2.51)	(2.00)
Initial Group$_{i,t}$	-13.72***	-13.74***	-13.68***	-12.48***	-8.67*
	(3.84)	(3.83)	(3.84)	(3.83)	(4.73)
Contribution$_{i,t}$ *- Group Contribution*$_{g[-i],t}$	3.04***	3.04***	3.03***	2.99***	1.77***
	(0.61)	(0.60)	(0.60)	(0.61)	(0.63)
abs(Group Size$_{g,t}$ *- 5)*	1.40	1.45	1.45	1.47	-0.84
	(0.92)	(0.92)	(0.92)	(0.95)	(2.19)
Vote Different from the Group$_{i,t-1}$		6.63**			
		(2.60)			
Vote Different from the Group$_{i,t-1}$ *under Rule-Set 1*			4.70	4.52	
			(4.19)	(4.18)	
Voted Against the Group$_{i,t-1}$ *under Rule-Set 2*			7.85**	8.06**	
			(3.41)	(3.42)	
Pre-Voting Period$_t$ *- Initial Group*$_{i,t}$				-1.30	
				(1.63)	
Pre-Voting Period$_t$ *- Outside Initial Group*$_{i,t}$				5.83**	
				(2.68)	
Observations	7,250	7,250	7,250	7,250	7,250

Notes. The table replicates the analyses in Table 4, reporting marginal effects of the covariates on the probability of moving (in percentage points), using Fernández-Val and Vella (2011)'s two-step (control function) approach for non-linear panel data models. Columns (1)-(4) are fitted to data from the *Voting-Moving* treatment, while column (5) uses data from the *No-Voting-Moving* condition. These specifications treat *Is Punishing*$_{g,t}$ (in the *Voting-Moving* treatment), *Initial Group*$_{i,t}$ and *abs(Group Size*$_{g,t}$ *- 5)* as endogenous, and allow for dynamic feedback from the dependent variable to the predictors. The control variables are the (generalized) residuals from the reduced-form models for the endogenous variables, specified as functions of pre-determined regressors - their own lagged values, the lagged average contribution of other members of i's group, the lagged values of $R_h(n)$, indicators capturing whether subjects' decision in the most recent voting round coincided with their/the other group's institutional choice (in columns 1-4), dummies for periods right before and after a voting round (in columns 1-4), and the lag of *Is Punishing*$_{g,t}$ (in column 5) - along with subject fixed-effects. All specifications in the table include subject fixed-effects; although time dummies (or time trends) are not covered by Fernández-Val and Vella (2011)' regularity conditions, adding them does not alter the results. Standard errors are reported in parentheses; we combined split-panel jackknife estimation and (panel) bootstrapping to correct for incidental parameter bias and adjust for generated regressors (Dhaene and Jochmans, 2015). Significance levels: *** at 1%, ** at 5%, * at 10%. The first-step (reduced-form) estimates are available from the authors upon request.

Table A.7
Parameter Estimates - Probit Panel Models for Voting

	(1)	(2)	(3)
Intercept	-2.11***	-2.83***	-2.05***
	(0.40)	(0.40)	(0.35)
$Is\ Punishing_{g,t-1}$	0.24*	0.21	-0.15
	(0.13)	(0.13)	(0.16)
$Initial\ Group_{i,t}$	0.43***	0.42***	0.31**
	(0.16)	(0.16)	(0.15)
$Contribution_{i,t}$ - $Group\ Contribution_{g[-i],t}$	-0.08	-0.08	-0.11*
	(0.06)	(0.06)	(0.07)
$abs(Group\ Size_{g,t}$ - 5)	0.10**	0.09*	0.11*
	(0.05)	(0.05)	(0.05)
Time Trend		0.04***	0.03***
		(0.01)	(0.01)
$Vote_{i,t-1}$			1.01***
			(0.15)
Observations	1,250	1,250	1,250
Log likelihood	-612.38	-605.16	-457.79

Notes. The table reports "raw" parameter estimates for the static (columns 1-2) and dynamic (column 3) panel probit models used to compute the marginal effects reported in Table 5. All specifications include subject, period, group, and session random effects; columns (2)-(3) also include a linear time trend. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table A.8
Parameter Estimates
Fixed-Effects Probit Panel Models for Voting

	(1)	(2)	(3)
$Is\ Punishing_{g,t-1}$	0.47***	0.44***	-0.07
	(0.14)	(0.14)	(0.20)
$Initial\ Group_{i,t}$	0.41**	0.38**	0.79***
	(0.18)	(0.18)	(0.22)
$Contribution_{i,t}$ - $Group\ Contribution_{g[-i],t}$	-0.05	-0.05	0.07
	(0.07)	(0.07)	(0.09)
$abs(Group\ Size_{g,t}$ - $5)$	0.00	-0.01	0.20***
	(0.06)	(0.06)	(0.07)
Time Trend		0.03***	-0.01
		(0.01)	(0.01)
$Vote_{i,t-1}$			12.32***
			(0.18)
Observations	665	665	420
Log likelihood	-349.14	-349.97	-230.42

Notes. The table reports parameter estimates from bias-corrected fixed-effects panel probit models for $Vote_{i,t}$ using Fernández-Val and Weidner (2016)'s method. The bias correction is obtained from jackknife estimates; using an analytical correction yields virtually identical results. Note that Fernández-Val and Weidner (2016)'s method relies on large-T (and large-N) asymptotics, and thus is not particularly well suited for the analysis of data from only five voting periods. Differences in the number of observations vis-á-vis Tables 5 and A.7 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 fixed-effects. Column (1) also includes period fixed-effects, which must be dropped in columns (2) and (3) due to the addition of a *Time Trend* to the model specification. Standard errors are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%.

Table A.9
Determinants of Individual Voting Decisions
Accounting for Covariate Endogeneity Using
Giles and Murtazashvili (2013)'s Control Function Approach

	(1)	(2)	(3)
$Is\ Punishing_{g,t-1}$	13.94***	12.81***	15.49***
	(4.00)	(3.90)	(5.09)
$Initial\ Group_{i,t}$	11.00***	10.46***	8.86*
	(3.60)	(3.44)	(4.88)
$Contribution_{i,t}$ - $Group\ Contribution_{g[-i],t}$	0.58	0.55	0.27
	(1.65)	(1.62)	(1.91)
$abs(Group\ Size_{g,t}$ - $5)$	3.10***	2.99***	3.10***
	(0.91)	(0.87)	(0.92)
Time Trend		4.64***	1.19
		(0.73)	(0.96)
$Vote_{i,t-1}$			43.60***
			(4.60)
Observations	1,250	1,250	1,000

Notes. The table replicates the analyses in Table 5, reporting marginal effects of the covariates on the probability of voting for institutions (in percentage points), using a control function approach for random effects panel probit models (Giles and Murtazashvili, 2013). Hausman tests justify the need to control for the endogeneity of $Initial\ Group_{i,t}$ and $abs(Group\ Size_{g,t}$ - $5)$ (p-values<0.01 across all specifications). We used as instruments the fourth and fifth lags of the average contributions of other members of i's group (Smith, 2013), an indicator for subjects who moved in the period before each voting round, and dummies for subjects who contributed more than 35 tokens in the first stage of the experiment - as a proxy for their "behavioral type" (Gürerk, Irlenbusch and Rockenbach, 2006). Diagnostic tests indicate that the instruments are valid and not weak (the first-stage F-statistics are always higher than 10) (Stock and Yogo, 2005; Wooldridge, 2010). As required by Giles and Murtazashvili (2013)'s method, we fitted linear first-stage regressions for $Initial\ Group_{i,t}$, which work well in practice for dealing with binary endogenous covariates (Angrist, 2001). As an alternative, we used the generalized residuals of probit models as control variables (Fernández-Val and Vella, 2011) and applied special regressor estimators (Dong and Lewbel, 2015), which are better suited for dealing with non-continuously distributed endogenous variables; the results are qualitatively similar. Bootstrapped standard errors adjusting for the first-stage estimation (Giles and Murtazashvili, 2013) are reported in parentheses. Significance levels: *** at 1%, ** at 5%, * at 10%. The first-stage estimates are available from the authors upon request.