Exp Econ (2015) 18:522–541
DOI 10.1007/s10683-014-9418-8

ORIGINAL PAPER

Voting with hands and feet: the requirements for optimal group formation

Andrea Robbett

Received: 13 September 2013 / Revised: 18 September 2014 / Accepted: 20 September 2014 /
Published online: 1 October 2014

Abstract This paper studies the dynamics by which individuals with heterogeneous preferences partition themselves into groups. A novel experimental environment is developed to capture the tension between increasing returns to group size and attaining a group policy closest to an ideal point. Subjects can move freely between locations, with group policy either fixed by location or determined by member vote. A primary goal is to assess which of two stability concepts common to the group formation literature predicts which groups agents sort into. The same set of *Nash stable* partitions exist in each condition, with the efficient, *strong Nash stable* state requiring subjects to form heterogeneous groups and compromise on policy. I find that subjects who are only able to move between locations with fixed policies always over-segregate, rather than build efficient heterogeneous groups. When mobility is combined with the ability to vote on local policy, most subjects reach the efficient partition. This shows outcomes cannot be determined by considering the existence of stable states alone and that consideration must also be given to subtle aspects of the system dynamics. Further, it suggests that experiments may play an important role in understanding these group formation dynamics.

Keywords Group formation · Heterogeneous preferences · Experimental economics · Tiebout · Experimental political science

JEL Classification C92 · C72 · C73

A. Robbett (✉)
Department of Economics, Middlebury College, Middlebury, VT 05753, USA
e-mail: arobbett@middlebury.edu

1 Introduction

Populations with diverse preferences must often form and sustain heterogeneous groups to take advantage of increasing returns to scale and capture gains from pooling resources. For instance, individuals with differing ideologies often join together to form a single political party or coalition, and neighboring towns often build single shared recreation facilities even though residents disagree on its ideal location. This paper uses laboratory experiments to study how agents partition themselves into groups when there are strong benefits to joining with others whose preferences for a group policy are not perfectly aligned with their own.

Overall, the experimental results suggest that the institution determining how group policies are chosen can greatly affect whether agents reach optimal partitions. When group policies are fixed, such that individuals can vote only with their feet by moving between locations, subjects fully segregate by preference type. However, when mobility is combined with voting, such that individuals may influence the policy of the group they join, most subjects succeed in forming groups of the optimal membership composition.

Two stability concepts are commonly considered side by side in the theoretical literature concerning the partitioning of agents across communities, clubs, or other groups. The concepts differ in whether a stable partition must be immune only to unilateral deviations or also to coalitional deviations. The first corresponds to the Nash equilibrium: a partition of agents is considered Nash stable if no agent can gain by unilaterally moving to a different group. Many Nash stable partitions typically exist and they are generally inefficient. The second, stronger stability concept corresponds to the strong Nash equilibrium (Aumann 1959) and requires that there is no set of individuals who could each do better by collectively relocating.[1] Such equilibria are Pareto efficient (though often nonexistent). If participants in a dynamic game myopically best respond to the previous state, then all Nash stable partitions will be absorbing. In other words, agents will be trapped in the first Nash stable state that they reach, even if more efficient outcomes could be possible with coordinated movement. It is unclear, however, whether agents will in fact myopically best respond in a dynamic group formation game, and experimental tests are thus necessary to determine which is the more appropriate stability concept. Will participants always remain in the first Nash stable partition they reach? Or will the system instead reach the more efficient strong Nash stable outcome?

The goals of this paper are: first, to provide a preliminary assessment of whether inefficient Nash stable states are, in fact, absorbing in a dynamic group formation game, or whether agents tend to reach efficient stable states when such exist; and second, to consider whether the means of establishing group policies determines into which stable partition agents sort themselves.

[1] This dichotomy originates with Tiebout (1956)'s canonical paper on free mobility and local public good provision, which seemingly refers to both types of deviations without distinction. Within these broad classifications, the definitions vary as well, both with respect to whether agents can relocate only to extant groups, or are able to establish groups of their own, and whether agents can coalitionally relocate only alongside those in their previous group or they are able to coordinate with any agents in the population.

This paper considers a simple experimental environment in which subjects can move freely between groups with various local policies. There are several available locations that remain fixed for the duration of the experiment and subjects play a 20-period dynamic game. In each period, subjects simultaneously choose their location. They receive a payoff based on the number of other subjects who chose the same location and the group's policy on some unidimensional issue. The policy is simply a number in the [0,1] interval, and subjects are assigned symmetric, single-peaked preferences over the interval at the start of the experiment. This policy can be interpreted as any outcome that applies indiscriminately to all group members, such as the local tax rate in a community, the platform of a political party or organization, the location of a club facility, or the type of good consumed within a club. The agents thus face a trade-off between being in a group of optimal size versus being in a group where the policy is closest to their ideal. The groups do not experience any congestion as membership grows, and so the optimal group size in this set-up is simply equal to the entire population. However, when agents have sufficiently divergent preferences and all members of a group are bound by the same local policy, they may receive higher payoffs by sorting into smaller groups with policies closer to their ideal.

Two conditions are conducted, which differ in how group policies are chosen. In the *Fixed Policy* sessions, each location is associated with a fixed, posted policy, which all group members experience in each period that they are in the location. In the *Voting* sessions, the policy is chosen in each period by member vote. In the environment considered in this experiment, all Nash stable partitions will be sorted, in the sense that agents of similar types will locate in the same group. In other words, if one considers the range of ideal points represented in each group, these ranges will not overlap across groups. However, the specification of these ranges may be inefficient, such that some Nash stable partitions are Pareto dominated. In these experiments, the same set of Nash stable partitions exists in both the Fixed Policy and the Voting conditions. Additionally, the same unique strong Nash stable partition exists in both conditions and occurs when two groups form, each comprised of those whose ideal points fall within one-half of the [0,1] interval. In other words, the optimal outcome for the population requires the formation of groups with heterogeneous membership and an intermediate, compromise policy, but there exist other, less efficient, equilibrium partitions in which more groups form.

Overall, I find that most subjects sort into a Nash stable partition. However, *which* partition they reach depends on how group policy is determined. I find that subjects who can vote only with their feet for group policies fail to partition themselves optimally. The subjects in the Fixed Policy sessions never succeed in reaching the optimal partition of two heterogeneous groups, and instead nearly all fully segregate into four homogeneous groups. This suggests that the existence of locations with fixed, posted policies facilitates the rapid sorting of agents by type into homogeneous groups, such as Tiebout envisioned, but may inhibit the formation of heterogeneous groups with compromise policies. Thus, fixing local policies may enable agents to successfully partition only when homogeneous groups are optimal, and may otherwise lead to over-segregation.

In stark contrast, subjects who are able to vote on their local policies, as well as move freely between groups, typically succeed in forming groups of optimal size and membership composition: The majority of subjects in the Voting sessions reach the strong Nash stable outcome by the end of the session. The prevalence of optimal, heterogeneous groups in the Voting condition is due both to the ability of larger groups to persist by internally adjusting local policy to changes in their membership composition and to the ability of subjects to merge pre-existing groups by implementing compromise policies. This suggests that the ability of group members to influence local policy, without needing to relocate, may not only be necessary for assuring that a group attains its optimal policy for a given membership composition once it has already sorted, but is also necessary for the population to reach the optimal formation.

The difference in outcomes between institutions occurs despite the equivalence in the set of Nash stable and strong Nash stable partitions under each. This suggests that how the population will partition cannot be assessed by solely considering the existence of stable states and that consideration must also be given to the system dynamics. In this case, it is the determination of local policy that alters these dynamics. Overall, subjects exit groups when they could have received higher payoffs elsewhere in the previous period, and unstable partitions rarely persist. Deviation from one of the inefficient Nash stable partitions is uncommon, and equally rare in both conditions. However, when such deviations do occur, the population is far more likely to transition to the efficient partition when the subjects are able to vote on local policy.

I find that participants in the Voting condition respond less to the current size and policy of groups than participants in the Fixed Policy condition do, and more to the presence of likeminded types, with whom they might influence local policy. At the group level, groups in the Voting condition are both more likely to grow in population and less likely to shrink than groups in the Fixed Policy condition, controlling for group features. The treatment differences in both the persistence of larger groups and the likelihood of pre-existing groups to merge can be directly traced to the number of moves required to initiate a policy change. In the Fixed Policy condition, the only means by which subjects can build compromise groups or alter the policy within a pre-existing group is to move to a new location and hope to attract members. The system is thus more likely to become "stuck" in an inefficient partition, as even when subjects deviate from an inefficient stable partition the system quickly returns to the same partition. In the Voting condition, less movement is required to reach the same outcomes, as groups can internally adjust their policy as their membership changes without the need to form new groups.

The rest of the paper is organized as follows: Sect. 2 reviews the related experimental and theoretical literature; Sect. 3 describes the experimental environment, procedure, and predictions; Sect. 4 presents the final group outcomes; Sect. 5 considers the dynamics that lead to these outcomes; and Sect. 6 concludes.

2 Related literature

The theoretical literature on endogenous group formation began with work in both local public finance (Tiebout 1956) and club formation (Buchanan 1965), though in neither original model does efficiency necessitate agents of different preference types joining together to form a stable group. In Buchanan's original model, all agents were identical and the balance of positive and negative externalities imposed by each implied a common and finite optimal group size. Tiebout, in contrast, considered a population with heterogeneous public good preferences, but also assumed an optimal group size that was small relative to the population, such that agents could select a community offering an expenditure package that precisely matched their preferences without sacrificing group size returns. Since agents could sort into optimally-sized, homogeneous communities, compromise was unnecessary and local governance irrelevant.

Beginning with Westhoff (1977), both the local public finance literature and club formation literature have explored the tension between increasing returns to group size and preference heterogeneity. A common conclusion is that, if there is a single parameter by which agents' preferences can be ordered along a single dimension (as is the case with the preferences assigned to subjects in the experiment in this paper), then an efficient, stable partition exists in which agents of similar preferences cluster together.[2] However, there has been little work on the dynamic aspect of group formation, and the conditions under which a population may in fact reach this efficient partition remain largely unexplored. One exception is Arnold and Wooders (2005), which presents a dynamic group formation model in which agents who are myopically best-responding to the previous partition converge to a Nash equilibrium, and agents who are able to communicate converge to the strong Nash stable equilibrium if such a partition exists. In the absence of communication, all stable partitions are absorbing, and so the system will end up in the first stable partition reached.

Experimental work has yet to directly address these dynamics or equilibrium concepts and there have been few experimental studies considering optimal partitioning of agents into groups. An over-arching conclusion of these experiments is that social preferences may inhibit the formation of optimally sized groups. Several experiments that allow members to control the size of their group in a pure public goods game, in which additional members are weakly beneficial, have found that subjects will often choose to expel, or approve the exit of, members from their group, or deny entry into their group.[3] Similarly, experiments have shown that

[2] Westhoff (1977) first formalized Tiebout's model, while removing congestion and incorporating majority rule voting on local tax rates. He proved existence of a stable partition of agents into several communities where the median voter's will was enacted, and that each community in this partition represented an interval of agents. Greenberg and Weber (1986) assume that agents' preferences can be ranked by a unidimensional parameter and demonstrate existence of an equilibrium partition immune to coalition deviations by secession. Demange and Henriet (1991) incorporate a similar assumption in a market for a differentiated consumption good with free-entry and demonstrate that an optimal, stable configuration of consumers across firms exists. Demange (2005) provides a more thorough overview of this tension between increasing returns to group size and preference heterogeneity.

[3] Cinyabuguma et al. (2005) and Maier-Rigaud et al. (2005) have found that expulsion is used frequently in public goods games. However, since the threat of expulsion also increases cooperation, subjects tend to earn more when expulsion is available. In contrast, Ahn et al. (2008) found that allowing subjects to control group size can suppress earnings in a pure public goods game. While restricted entry enables groups to sustain higher contribution rates, groups of cooperative subjects tend to earn less by being overly discerning in whom they allow to enter.

cooperative subjects will themselves exit efficient, strong Nash stable partitions to escape less cooperative subjects.[4] Crosson et al. (2004) directly test subjects' ability to partition themselves into optimally sized groups by distributing complementary resources (playing cards) and allowing them to form groups to produce a club good (poker hand), which is then divided among the members. Though only three players were necessary to produce the good, few groups restricted their membership to the optimal size, choosing to instead allow additional members to join.

Moving beyond the question of optimal group size, the experiment reported in Robbett (2014) is more directly related to whether individuals with different preferences sort themselves into groups according to these preferences. That paper tests Tiebout's premise that mobility can lead to efficient public good provision, by studying an environment in which subjects receive either very high or very low returns from public good provision and can move between communities with local tax policies. When all residents of a community are required to pay equal taxes, subjects in that experiment do sort by type into two homogeneous communities. Despite sorting, however, they often fail to achieve the level of public good provision best suited to their type. When participants can vote, inefficient equilibria are eliminated and the homogeneous communities enact the optimal public good provision for their residents.

This paper instead focuses on the partitioning of subjects into groups. As described in Sect. 3, a range of preference types is generated such that multiple stable, "sorted" partitions are possible in equilibrium. Further, the set of stable partitions is identical across the two treatments. The results reported in this paper thus offer a surprising complement to the previous work: granting participants the ability to vote can lead to more efficient outcomes, not only by eliminating less efficient equilibria in which groups enact the wrong policy for their population, but also by promoting the optimal partitioning *into* groups, which would not be predicted by simply comparing the set of stable partitions across treatments.

3 Experimental design

3.1 Setting

There is a finite set of agents $N = \{1, \ldots, n\}$ that differ only in a unidimensional preference parameter $\alpha \in [0, 1]$. Let $X = \{S_g\}_{g=1}^{k}$ be a partition of the set N into k groups such that $\cup_{g=1}^{k} S_g = N$ and, for all $g, h \in \{1, \ldots, k\}$, $S_g \cap S_h = \emptyset$.

A *state* $(X, \mathbf{p})$ is a partition X of the agents into k groups and a k-tuple of policies $\mathbf{p} = (p^1, \ldots, p^k)$, where $p^g \in [0, 1]$ for all g. In other words, each agent is a member of exactly one group and each group is associated with a local policy that governs all members.

Let $S_X(i)$ be the group to which i belongs in partition X and $X(i)$ refer to its index. Agents have preferences over the size of their group and over the policy space. For a

[4] Robbett (2010, 2014) each find that frequent movement leads to suppressed earnings in linear and non-linear pure public goods games with free mobility.

given group size, each agent i has single-peaked preferences over p in [0,1], with payoffs maximized at α^i. Additionally, payoffs are increasing in group size.

In particular, agent i receives a payoff from being a member of group $S_X(i)$:

$$\pi^i(X, \mathbf{p}) = |S_X(i)| - \gamma(\alpha^i - \mathrm{p}^{X(i)})^2 \tag{1}$$

where $\gamma > 0$ and $|S_X(i)|$ is the size of the group to which agent i belongs under the partition. This is a simple representation of the ubiquitous trade-off between being in an optimally-sized group (which, in the congestion-free environment considered here, is equal to n) and being in a group where the local policy best matches one's ideal. This payoff function can be thought of as representing any environment in which all agents are bound by a single local policy over which they have single-peaked preferences. For instance, away from the extremes, this function is very similar to the payoffs that agents with Cobb-Douglas preferences for public good provision and private consumption have over a local tax rate and community size, though with the feature that agents of all values of α are equally willing to compromise.

3.2 Experimental procedure

The experiment was conducted at the Harvard Decision Science Laboratory in Cambridge, Massachusetts. Overall, 120 undergraduate and graduate students participated. While groups of 16 or 24 students participated in each session, the subjects were further divided into "populations" of eight people. In the experiment, they interacted only with the participants in their own population. The subjects participated using the experimental software z-Tree (Fischbacher 2007) and were unaware which of the other people in the room they were interacting with.

Subjects were randomly assigned different preferences over local policies, implemented through the value α in the payoff function in Eq. (1). Within each population of eight subjects, exactly two people were assigned to each of four different types: $\alpha = .15$, $\alpha = .35$, $\alpha = .65$, and $\alpha = .85$. The set of α values was selected in order to generate several stable "sorted" partitions, which could be Pareto ranked and comparable across the institutions for determining local policies. The subjects played a repeated game that lasted twenty periods. There were seven available locations, which remained fixed for the duration of the experiment and were labeled "Group 1" through "Group 7." At the beginning of each period, the subjects simultaneously chose a location. While making this choice, they were able to observe the number of members and the policy in each of the seven locations in each of the previous three periods.[5] However, they did not receive information on the identity of these subjects or their ideal policies. After choosing their location, the subjects were told the number of others who chose the same location, and a local policy was implemented for everyone in the group. Finally, they received the payoff given in Eq. (1), where γ, the parameter specifying the trade-off between group size

[5] In the first three periods, subjects viewed the location features for the duration of the experiment thus far.

and policy, was set equal to 60, for reasons described in the final paragraph of the following section.

Subjects were in one of two conditions, and how their group policy was determined depended on the condition of their experimental session. The conditions follow the two institutions most commonly considered in the group formation literature. In the *Fixed Policy* sessions, each location was associated with a different fixed, posted policy. Among these locations were those offering the ideal policies for each of the four types as well as the three "compromise" policies that were midway between each of the types' ideals (.25, .5, and .75). Both the number of locations and the specific set policies were selected to generate an equivalent set of equilibrium partitions across conditions, as described the subsequent section. In the *Voting* sessions, the groups' current members voted on their policy in each period, with the median voter's preference implemented.[6] Seven populations were run for the Fixed Policy condition (n = 56) and eight populations were run for the Voting condition (n = 64). The experimental procedure for each period, under each condition, is summarized in Fig. 1.

In the first period, the subjects were initially assigned to a single group of all eight participants. In the Fixed Policy sessions, the policy in this initial location was 0.5, which is the policy that would be enacted in the Voting sessions if all subjects voted for their ideal policy. In subsequent periods, subjects were free to choose whichever group they wished. However moving—choosing a different location than in the previous period—carried a small moving cost of 0.3 experimental units.

3.3 Stability predictions

A partition of subjects is *Nash stable* if no subject can receive higher payoffs by unilaterally relocating to a different group. A partition is *strong Nash stable* if there is no set of agents who can each receive higher payoffs by collectively relocating.[7]

There exist the same Nash stable partitions in both the Fixed Policy and the Voting sessions. These partitions are depicted graphically in Fig. 2. Each thick horizontal line represents a populated location, and each box stacked above depicts an agent at that location. The number in the box gives the individual's assigned α value. The numerical range beneath each location shows the range of policies in each group for which the partition is Nash stable (with the number above the arrow indicating the minimum distance between adjacent policies necessary for stability in

[6] In the case that the group had an even-numbered population, the median policy was equal to the mean of the two middle votes.

[7] While Nash stability requires that an outcome be immune to unilateral deviations, strong Nash stability requires that the outcome be immune to deviations by *any subset* of agents. Another common stability concept that considers (nested) coalitional deviations is the *coalition-proof Nash equilibrium* (Bernheim et al. 1987). It is less stringent than the strong Nash equilibrium, as it does not require that partitions be immune to all coalition deviations, but only those that are "self-enforcing," i.e. from which no subset of the coalition would wish to further deviate at any step. In the experimental set-up of this paper, the unique strong Nash partition is also the unique coalition-proof partition.

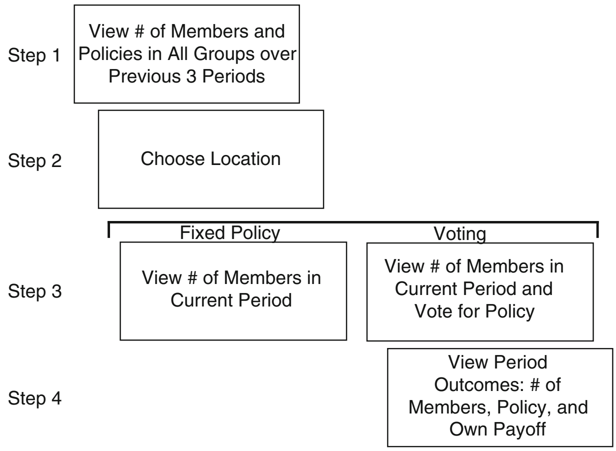

Fig. 1 Experimental procedure in each period

the most extreme case). Finally, the figure also shows the maximum payoff that the population of participants could achieve by forming this partition.[8]

The first Nash stable partition is a *Segregated* partition in which each of the types are in their own separate group, with a policy close to their ideal. The second, *Center Pooled*, Nash stable partition has the two extreme types segregated in their own groups while the two moderate types form a single group with a policy close to the center. The third, *Pareto Dominant* partition, has agents forming two large groups with those whose preferences are similar, with a policy within or very close to the range of policies represented in the group. Aggregate payoffs are maximized when agents form the Pareto Dominant partition. There are two other asymmetric stable partitions that are combinations of the Segregated and the Pareto Dominant partitions. In one (pictured here), only those whose ideal points lie within the lower half of the [0,1] interval join together while those whose ideal points lie within the upper half segregate; the other has those whose ideal points lie in the *upper half* of the [0,1] interval pooling together and the others segregating. Note that the locations' policies in the Fixed Policy condition enable each of these partitions to be reached and fall at the midpoint of each of the ranges listed in the figure. This is essential for guaranteeing that the set of stable partitions is the same across conditions.

Finally, there is a unique strong Nash stable partition, which is the same under both the Fixed Policy and Voting conditions, and is identical to the Pareto dominant Partition, but with a slightly narrower range of supported policies.

Note that, though the set of Nash stable partitions is identical under both the Fixed Policy and Voting conditions for these parameter values, this need not hold generally. For instance, consider a similar environment with eight agents for whom $\gamma = 45$ rather than 60, and α values of .05, .15, .2, .35, .65, .8, .85, and .95,

[8] While the exact payoff for each agent depends on the precise policy p, aggregate payoffs are maximized in each case when the group experiences the p at the midpoint of the range displayed. The full derivation is provided in the online appendix.

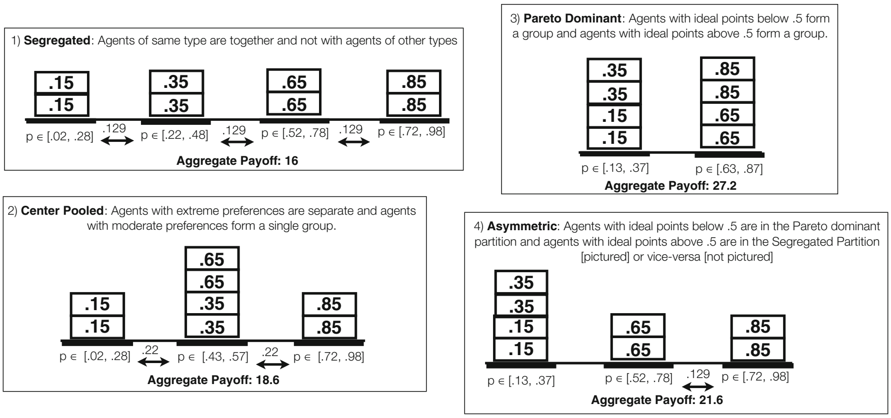

Fig. 2 Forms of nash stable partitions

respectively. The state in which the three agents with the lowest ideal points are in a group with policy .05, the three agents with the highest ideal points are in a group with policy .95, and the two most moderate agents are in a group with policy .5 is Nash stable under a Fixed Policy institution (with an infinite array of available locations). In fact, it is the state at which the system would arrive if all agents began in a single group with a center policy and myopically best-responded. However, if the agents in this partition voted, then the policies in the extreme groups would move closer to the center, and the two agents with moderate preferences would do better by joining them. Therefore, this is an example of a partition that is Nash stable in the Fixed Policy condition but not in the Voting condition.

4 Results: final outcomes

This section will focus on the outcomes that participants reach by the end of the experimental session; Sect. 5 will focus on the dynamics leading to these outcomes.

4.1 Efficiency convergence

We first examine the efficiency of the outcomes to which the subjects converge in each condition, and find that subjects receive significantly higher payoffs when they are able to vote on their local policies. Figure 3 shows the average aggregate (population-wide) payoffs over the final five periods of the twenty period experiment under the Fixed Policy and Voting conditions. The dashed line shows the payoff under the most efficient equilibrium outcome (also the strong Nash state) while the dotted line shows the payoff under the least efficient equilibrium outcome (the Segregated partition). I find that the aggregate payoffs are higher in the Voting condition than in the Fixed Policy condition and this difference is significant at any reasonable level.[9] Aggregate payoffs in the Voting condition are between those under the least efficient stable outcome and the most efficient stable outcome, while the aggregate payoffs in the Fixed Policy condition are slightly (though not significantly) less than those in the least efficient stable outcome. This suggests that the subjects in the Fixed Policy condition either are not reaching a stable partition or are at one of the least efficient partitions, with some deviations.

4.2 Group features

We next examine the source of this difference by considering the features of the groups in existence at the end of the experimental session in each condition. Overall, the subjects in the Voting condition tend to sort into much larger groups with local policies further from their ideal. Table 1 shows the frequency with which subjects end up groups of various sizes and policies during the final five periods of the experiment.

Comparing the group size across the two conditions, we see that subjects in the Voting condition form larger groups. Specifically, subjects in the Voting condition are most typically in a group of 4, while subjects in the Fixed Policy condition are in a group with only one other member in more than three-quarters of observations. The difference in group size between conditions is significant at the 1 % level.[10] Consistent with the prediction in the theoretical literature that individuals will form a "sorted" partition, subjects in both conditions are grouped with the one other individual who shares the same ideal point the vast majority of the time.[11] Subjects in the Voting condition usually also share a group with individuals who do not have the same preferences, while this happens only rarely in the Fixed Policy condition.[12]

By comparing the number of observations in each row across conditions in Table 1, we see that the subjects in the Fixed Policy condition attain policies that are closer to their ideal points.[13] In fact, subjects in the Fixed Policy condition experience their exact ideal policy most of the time (84 % of observations). This indicates that the lower efficiency of the Fixed Policy condition is not driven by failure to sort into a group with the optimal policy, but instead by the formation of smaller groups. In contrast, subjects in the Voting condition experience their ideal policy less than half of

[9] The average efficiency over the final five periods of one population of 8 subjects is taken as a single observation. The difference between conditions is significant at less than the 1 % level using a two-sample Wilcoxon rank-sum test (z = 2.664).

[10] Taking the average group size in a population over the final five periods as a single observation, the difference in group size is significantly higher in the Voting condition at the $p<.01$ level using a Wilcoxon rank-sum test (z = 3.30).

[11] Subjects are in the same group as the person who shares their preference type in 85 % of observations in the Fixed Policy condition and 94 % of observations in the Voting condition (over the final five periods).

[12] Subjects in the Voting condition are in a group with subjects of other types 61 % of the time in the Voting condition, compared to 16 % of the time in the Fixed Policy condition. The difference is significant at $p<.01$ level using a Wilcoxon rank-sum test (z = 3.309).

[13] In particular, subjects in the Fixed Policy condition experience policies that differ from their ideal by only .026, while the average difference in the Voting condition is .071. This difference is significant at $p<.01$ using a Wilcoxon rank-sum test and taking the population as the unit of observation (z = 2.78).

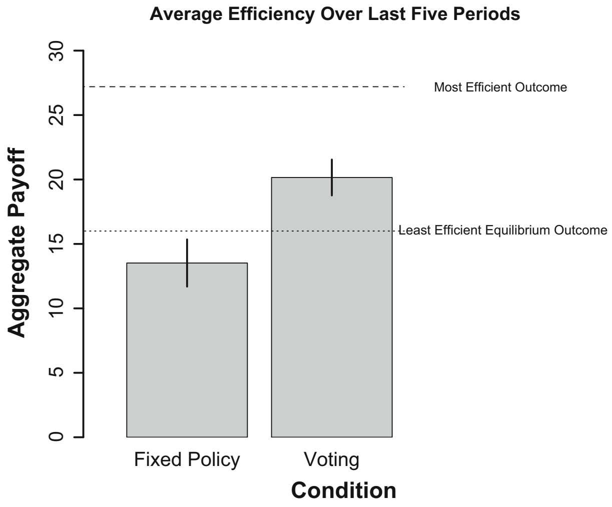

Fig. 3 The average efficiency over the final five periods (means taken at population-level)

Table 1 Group features in the Fixed Policy condition and Voting condition

	Fixed Policy					Voting					
	Group size					Group size					
	1	2	3	4	Total	1	2	3	4	5	Total
$\lvert$Policy − ideal point$\rvert$ = 0	17	200	12	6	235	9	82	8	4	0	103
0 < $\lvert$Policy − ideal point$\rvert$ ≤ .05	0	0	0	0	0	0	20	0	22	0	42
.05 < $\lvert$Policy − ideal point ≤ .1	2	0	0	0	2	0	2	3	80	0	85
.10 < $\lvert$Policy − ideal point$\rvert$ ≤ .15	5	13	8	0	26	3	0	0	44	2	49
.15 < $\lvert$Policy − ideal point$\rvert$ ≤ .2	1	1	4	0	6	0	6	2	24	1	33
.20 < $\lvert$Policy − ideal point$\rvert$	2	0	3	6	11	2	0	2	2	2	8
Total	27	214	27	12	280	14	110	15	176	5	320

This table reports one observation for each subject in each of the final five periods of the experiment. The columns show the size of the group that the subject is in and the rows show the absolute distance between the subject's ideal point and the group's current policy

the time (36 % of observations). This implies either that the subjects are compromising on policy or that they err in voting in their ideal policy. To consider this second explanation, we look at whether subjects experience a policy that is within the range of ideal points represented in the group. We find that this occurs upwards of 92 % of the time and is equally likely in both conditions. This suggests that deviations from the subjects' ideal policies in the Voting condition are driven primarily by ability to form groups with compromise policies, rather than by error.

4.3 Final partitions

Finally, we consider whether the partitions that subjects reach are consistent with the stability predictions that were introduced in Fig. 2. As indicated by Panel 4 in Fig. 2, it is possible for subjects who have ideal points on one side of the spectrum to be in a partition consistent with one classification (for instance, pooled together in the ideal configuration for these individuals), while those on the other side are in a partition consistent with a different classification (for instance, segregated) or with no stability prediction at all. Within a population, we will thus classify separately the partitions reached by the four subjects for whom $\alpha < .5$ and the four subjects for whom $\alpha > .5$. Figure 4 shows the partition classification over time for each experimental population. For each period, a marker appears indicating which of the stability classifications from Fig. 2, if any, describes the subjects' partition in that period. If the partition does not align perfectly with one of the stability predictions, no marker appears. The classification of those for whom $\alpha < .5$ is shown directly above the period number and the classification of those for whom $\alpha > .5$ is shown just to the right of it. Note that this classification is rather strict and if just one of the four subjects is out of position, or if the group's policy is out of the range listed in Fig. 2, then no marker will appear.

Remarkably, most subjects reach one of the predicted partitions: over the final five periods of the experiment, two-thirds of subjects in the Fixed Policy condition and four-fifths of subjects in the Voting condition are in a partition consistent with one of the stability predictions. Further, we see a sharp difference in *which* partition the subjects reach in the two conditions. In the Fixed Policy condition, the stable partition reached is nearly always the Segregated partition (Panel 1 in Fig. 2), whereas the modal partition reached by the Voting subjects is the efficient Pareto Dominant partition (Panel 3 in Fig. 2).

5 Results: overall dynamics

The previous section reported strong evidence that subjects who could vote for their local policy eventually sorted themselves into larger, heterogeneous groups and earned higher payoffs. In this section, we investigate the dynamics in order to uncover the mechanism that drives the differences in final outcomes.

To begin to understand the dynamics of group formation, it is important to first assess the factors that influence movement and whether Nash stable states are, in fact, absorbing. We can easily dismiss the premise that participants always myopically best respond to the previous partition. If this were the case, the individuals with the more extreme preferences ($\alpha = .15$ and $\alpha = .85$) would exit in their first opportunity and bring the system to the Center Pooled Nash stable partition (Panel 2 in Fig. 2) in the second period, where it would remain. However, as we saw in Fig. 4, the system is never in such a partition by the end of the experimental session.

We therefore turn to the question of which factors do influence movement. Table 2 provides summary statistics on the likelihood that an individual moves to a different

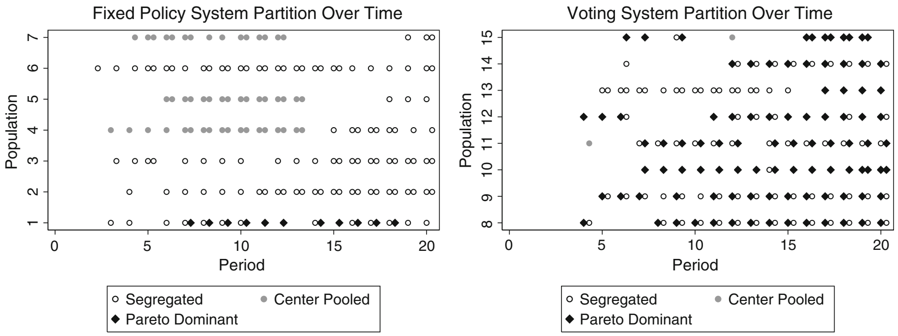

Fig. 4 Classifications of system partition over time, by experimental population

Table 2 Summary statistics: likelihood of movement in Fixed Policy and Voting conditions

Event	Likelihood in		Rank-sum test (z)
	Fixed Policy	Voting	
Moving in period 1	0.625	0.298	3.589***
Moving in period > 1	0.11	0.099	0.847
Remaining in initial group for all periods	0	0.219	−3.708***
Unilateral improving move available	0.162	0.176	−0.924
Moving when unilateral move available	0.597	0.482	2.272**
Moving when unilateral move not available	0.046	0.027	2.214**
Moving when unilateral move not available (period > 1)	0.035	0.024	1.376

* Significant at 10 % level; ** significant at 5 % level; *** significant at 1 % level

group in a given period. First, we see that individuals in the Fixed Policy condition are far more likely to exit the initial group in their very first opportunity. Following the first period, however, movement is rare in both conditions, with individuals moving in approximately one-tenth of their opportunities. The rarity of movement is not necessarily surprising given how often participants *do not* have a unilateral improving move available to them (nearly 85 % of the time in both conditions). Moreover, whether such an improving moving exists is highly predictive of whether the participant moves in the period. When a profitable unilateral move is available, subjects do frequently exit their group (60 % of the time in the Fixed Policy condition and 48 % of the time in the Voting condition). In sharp contrast, participants who do not have a unilateral improving move available remain in their group in over 95 % of opportunities. At the systemic level, Nash stable states persist into the next period 82 % of the time in the Fixed Policy condition and 89 % of the time in the Voting condition and there is no significant difference between the conditions.[14] Thus we find

[14] In the Voting condition, the strong Nash stable state persists with a similar likelihood; in the Fixed Policy condition, the system does not reach the strong Nash stable state.

that Nash stable states are not always absorbing (as assumed by Arnold and Wooders 2005 in the absence of communication), but they do persist with high likelihood.

To further investigate the decision to move to a different group, I estimate a series of linear probability models in which a dummy variable indicating whether the subject exits his group is regressed on explanatory features of the group in the previous period. These estimates are reported in Table 3. In each model, standard errors are clustered at the population level and the initial movement in the first period is excluded. First, we see that participants are less likely to exit larger groups and more likely to exit groups with policies further from their ideal points, both of which are highly significant (Column 1). Next, we again see that individuals are equally likely to exit a group in the Voting condition, even when controlling for these group features (Column 2). However, we do identify two differences between treatments. From the model that interacts a dummy for Voting condition with group features, we see that individuals in the Voting condition are *less* responsive to differences in group size and distance from ideal point than are individuals in the Fixed Policy condition (Column 3). The estimates reported in the two final columns only consider observations in which the participant is in a Nash stable or unstable state, respectively. We see that individuals in the Voting condition are less likely to move from unstable states, controlling for group features (Column 5). Thus we find that individuals in the Voting condition respond less to the observable features of the group, and that they are more willing to remain in their group when a unilateral improving move is available.

While the estimates reported in Table 3 provide insight into the subjects' decision to exit their group, to ensure that we have a complete picture of the participants' decisions, we also address the question of which groups the subjects choose to join. Table 4 thus presents estimates of a series of linear probability models with one observation for each group and each subject in each period. The dependent variable in these models is a dummy variable indicating whether the subject is in the group in the current period and the explanatory variables are features of the group from the previous period. We again find that participants respond strongly to both the population size and the difference between the group's policy and their own ideal point, but that participants in the Voting condition respond less to these features than their counterparts in the Fixed Policy sessions do.

Finally, we investigate the group-level dynamics by turning to the question of how group size is influenced by previous group features. The regression estimates reported in Table 5 take each populated group as a single observation and consider as the dependent variable changes in group size (Column 1) or, more specifically, whether the group shrinks (Column 2) or grows (Column 3) from one period to the next. We see that, controlling for previous size and members' preferences, voting has a positive effect on the change in group size from one period to the next. Groups are both more likely to grow and less likely to shrink when members can vote. This suggests two mechanisms by which subjects in the Voting condition succeed in forming larger, more efficient groups: Large groups are more likely to persist and smaller groups are more likely to expand or merge. Both of these mechanisms are strongly supported by direct inspection of the data and discussed in detail below.

Table 3 Moving decision

	(1)	(2)	(3)	(4) Stable	(5) Unstable
Group size	-0.0932^{***}	-0.0925^{***}	-0.137^{***}	-0.0147^{**}	-0.0679^{**}
	(0.00991)	(0.0109)	(0.0149)	(0.00513)	(0.0236)
| Policy − ideal point |	1.938^{***}	1.938^{***}	2.258^{***}	0.456^{***}	1.522^{***}
	(0.112)	(0.111)	(0.163)	(0.0821)	(0.147)
Voting condition		−0.00534	-0.143^{***}	−0.0109	-0.117^{***}
		(0.0149)	(0.0439)	(0.00847)	(0.0325)
Group size × voting			0.0633^{***}		
			(0.0173)		
| Policy − ideal point | × voting			-0.466^{**}		
			(0.194)		
Constant	0.237^{***}	0.238^{***}	0.327^{***}	0.0570^{***}	0.459^{***}
	(0.0262)	(0.0247)	(0.0326)	(0.0148)	(0.0414)
Observations	2,160	2,160	2,160	1,818	342
Pseudo R^2					

Standard errors in parentheses

Dependent variable is 1 if agent exits group. OLS regression reported

$^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$

First, the initial group is far less likely to deteriorate in the Voting condition, as indicated by Table 2. In the Fixed Policy condition, nearly all of those with extreme preferences and 40 % of those with moderate preferences exit the initial group in the first period. As this is not a myopic best-response for the moderates, this indicates that the existence of groups with the subjects' posted ideal policy is driving initial exit and segregation by providing a clear signal of which group to choose. In the Voting condition, less than half of those with extreme preferences and only 12.5 % of moderates exit initially. Unless this exit is perfectly symmetric, the distribution of preferences remaining within the group tips toward one side of the [0,1] interval: in all but one of Voting populations, the initial group implements a policy favoring those on one side of the spectrum by the second period, prompting those subjects to remain and the others to leave. In half of the Voting populations, all four of the subjects whose ideal points fall on that side of the interval remain in the group for the duration of the session. In contrast, no subjects in the Fixed Policy condition remain in the initial group by the end of the session.

However, this persistence of the initial group still does not account for the majority of four-member groups in existence at the end of the Voting sessions. Perhaps more important, as it does not depend on initial conditions, is the relative success of subjects in building up compromise groups after they have already scattered and segregated.

To illustrate how the process of building the larger compromise group differs between conditions, consider the situation in which the four participants with ideal points on one side of the spectrum have already segregated into two groups of two

Table 4 Group choice

	(1)	(2)	(3)	(4)
In group last period	0.878***		0.822***	0.802***
	(0.0110)		(0.0172)	(0.0226)
Previous | Policy − ideal point |		−0.559***	−0.108***	−0.168***
		(0.0351)	(0.0103)	(0.0162)
Previous group size		0.0976***	0.00994***	0.0152***
		(0.00332)	(0.00206)	(0.00281)
In group last period × voting				0.0248
				(0.0350)
Previous | Policy − ideal point | × voting				0.0736***
				(0.0192)
Previous group size × voting				−0.00611
				(0.00416)
Voting				−0.00646
				(0.00731)
Constant	0.0174***	0.232***	0.0529***	0.0602***
	(0.00158)	(0.00683)	(0.00337)	(0.00530)
Observations	15,120	15,120	15,120	15,120
Adjusted R^2	0.772	0.365	0.777	0.778

Standard errors in parentheses

Standard errors clustered at population-level. Does not include response to initial period.

$^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$

Table 5 Group dynamics

	(1) Δ size	(2) Shrinks	(3) Grows
Voting condition	0.0917***	−0.0415*	0.0466**
	(0.0355)	(0.0219)	(0.0209)
Group size	0.00188	−0.0693***	−0.0581***
	(0.0199)	(0.0123)	(0.0117)
Mean | Policy—ideal point |	−1.274***	0.352	−0.616**
	(0.440)	(0.272)	(0.260)
Max. | Policy—ideal point |	−1.053***	1.518***	0.443**
	(0.313)	(0.193)	(0.184)
Constant	0.0284	0.212***	0.218***
	(0.0437)	(0.0270)	(0.0258)
Observations	949	949	949

Standard errors in parentheses

Does not include initial period

$^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$

with policies equal to their ideal points. In this partition, each subject earns a per-period payoff of 2 units. First, we will consider how subjects in the Fixed Policy condition might attempt to transition from two segregated groups to one four-person group. One approach would be for a subject to deviate from the stable partition in favor of joining the group populated by the nearest type. This is illustrated in the far left panel of Fig. 5. In the figure, a white box indicates a subject's current position while a gray box indicates the subject's former position before moving along the arrow path. Here, one of the subjects for whom $\alpha = .15$ prompts the transition by joining the group where the policy is .35. At this point, the best response of his partner who shares his ideal point is to follow him into the group as well.[15] Thus, in just two moves, the subjects have transitioned to the four-person group, which provides higher mean payoffs than the two smaller groups did. However, now the two subjects who moved receive a payoff of only 1.6 each, instead of the payoff of 2 each that they would have received by remaining in the location associated with their ideal policy. Therefore it is not in a subject's interest to initiate this transition. Unsurprisingly, this is attempted only rarely—three times—and leads to the successful creation of a four-person group only once.

Alternatively, a subject in the Fixed Policy condition might attempt a transition to a four-person group with a compromise policy by moving into the empty location between the two ideal points. This scenario is depicted in the middle panel of Fig. 5. Again, his partner's best response is to follow him into the group. At this point, the other two subjects would be better off moving into the compromise group as well. This transition thus requires all four subjects to move locations and leads to increased earnings for each individual (per-period payoffs increase from 2 to 3.4). There are 10 observations in the data of this type of transition being attempted in the Fixed Policy sessions when the subjects were previously in a fully segregated partition. Though the subjects who initiate it remain in the compromise location for an average of 2.3 periods following their move, none succeed in attracting the three other subjects with similar preferences, and they eventually return to the segregated partition. This strategy is used only once in the Voting condition, and also fails.

However, when policies can be adjusted internally by group vote, fewer moves are required to form a larger compromise group. The scenario in the Voting condition is depicted in the right panel of Fig. 5. Initially, one subject must move to form a group of three, where he will receive a lower payoff. However, as soon as his partner joins him, they can influence the group's policy and move it closer to their own ideal, leading to an outcome similar to that depicted in the center panel. Thus only two moves are required to reach the strong Nash stable outcome, rather than four. This strategy is attempted in eight instances in the Voting condition, and the final subject nearly always follows his partner, completing the successful transition to the Pareto Dominant state.

[15] Recall that subjects incur a moving cost of .3 in a period in which they switch groups. Here and in the discussion to follow, a move is only considered a "best response" if the player's payoff increases in the immediate period even after accounting for the moving cost. The presence of the moving cost does not affect the best responses in any case. In the case considered here, by switching groups this subject improves his earnings by .3 in the period that he moves and .6 in each period after.

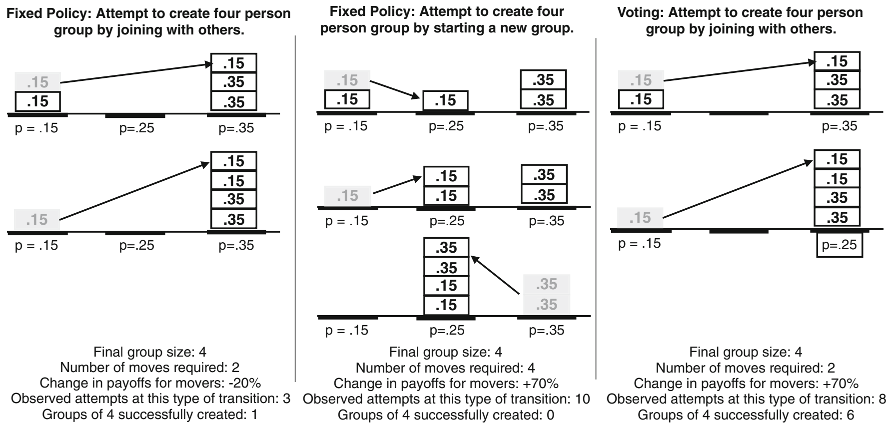

Fig. 5 Means by which subjects transition to larger groups and success rates. Each *box* refers to one subject and the *number within the box* refers to this individual's ideal point (α). A *white box* indicates a subject's current position. A *gray box* indicates the subject's former position before moving along the arrow path. The *number below the stack of boxes* indicates the group's current policy

6 Conclusion

The results reported in this paper suggest that the ability to vote with one's feet may not be sufficient for a population to reach an optimal partition when there are gains to be made from agents pooling resources with those who have similar preferences. While the existence of locations with fixed, posted policies enables the rapid sorting of subjects by type, the formation of homogeneous, segregated groups may not always be optimal. In that case, the saliency of locations offering a subject's ideal policy leads to initial over-movement and segregation, and the inability of newcomers to influence local policy deters the formation of compromise groups.

When subjects are able to vote on group policy, as well as move freely between groups, they most commonly succeed in forming larger heterogeneous groups of like-minded, though not identical, individuals and reaching optimal partitions. The ability of groups to internally adjust their policies in response to changes in membership composition enables subjects both to sustain larger groups and to expand pre-existing groups by implementing compromise policies. This ability to transition to the strong Nash partition without needing to establish new groups enables subjects to avoid being stuck in less efficient stable partitions.

These results suggest that the determination of local policy is critical for understanding whether diverse populations are likely to succeed in partitioning themselves optimally across groups. It appears that mobility in itself is not sufficient for reaching optimal partitions, and that additional mechanisms, such as a means of signaling or the ability to internally adjust local policies to the preferences of the current population, may also be necessary to enable groups to combine their resources efficiently. Furthermore, this difference between institutions occurs despite the equivalence in the set of equilibria under each. This suggests that institutional and environmental features alter the dynamics in ways that cannot be fully captured by considering stability concepts alone, and that additional laboratory experiments are necessary for understanding how agents sort themselves by preferences. For instance, one useful extension would be to determine whether allowing participants to communicate reduces the efficiency gap between the Voting and Fixed Policy institutions, by enabling all agents to reach the efficient outcome (as in Arnold and Wooders 2005).

Acknowledgments I thank Charles Plott, Leeat Yariv, Peter Matthews, Guillaume Fréchette, Rod Kiewiet, John Ledyard, Jean-Laurent Rosenthal, and the Harvard Decision Science Laboratory. I am also grateful to two anonymous referees and the editor, David Cooper, for their many helpful suggestions. Funding provided by the Caltech Laboratory for Experimental Economics and Political Science.

References

Ahn, T. K., Isaac, R. M., & Salmon, T. (2008). Endogenous group formation. *Journal of Public Economic Theory*, *10*(2), 171–194.

Arnold, T., & Wooders, M. (2005). Dynamic club formation with coordination. Working Paper No. 05–W22, Department of Economics, Vanderbilt University.

Aumann, R. J. (1959). Acceptable points in general cooperative n-person games. In *Contributions to the theory of games IV*. Princeton University press.

Bernheim, B. D., Peleg, B., & Whinston, M. D. (1987). Coalition-proof nash equilibria. I. Concepts. *Journal of Economic Theory*, *42*, 1–12.

Buchanan, J. M. (1965). An economic theory of clubs. *Economica, New Series*, *32*(125), 1–14.

Cinyabuguma, M., Page, T., & Putterman, L. (2005). Cooperation under the threat of expulsion in a public goods experiment. *Journal of Public Economics*, *89*(8), 1421–1435.

Crosson, S., Orbell, J., & Arrow, H. (2004). Social poker: A laboratory test of predictions from club theory. *Rationality and Society*, *16*, 225–248.

Demange, G. (2005). Group formation: The interaction of increasing returns and preferences diversity. In G. Demange & M. Wooders (Eds.), *Group formation in economics: Networks, clubs, and coalitions*. Cambridge, MA: Cambridge University Press.

Demange, G., & Henriet, D. (1991). Sustainable oligopolies. *Journal of Economic Theory*, *54*(2), 417–428.

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

Greenberg, J., & Weber, S. (1986). Strong Tiebout equilibrium under restricted preferences domain. *Journal of Economic Theory*, *38*, 101–117.

Maier-Rigaud, F. P., Martinsson, P., & Staffiero, G. (2005). Ostracism and the provision of a public good: Experimental evidence. *Preprints of the Max Planck Institute for Research on Collective Goods*, *2005*, 24.

Robbett, A. (2010). Community dynamics in the lab: Public good provision, congestion, and local instability. Caltech HSS Working Paper, 1339.

Robbett, A. (2014). Local institutions and the dynamics of community sorting. *American Economic Journal: Microeconomics*, *6*(3), 136–156.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, *64*, 416–424.

Westhoff, F. (1977). Existence of equilibria in economies with a local public good. *Journal of Economic Theory*, *14*, 84–112.