Endogenous Group Formation and Public Goods Provision: Exclusion, Exit, Mergers, and Redemption

Gary Charness & Chun-Lei Yang*

September 29, 2008

Abstract: We test a mechanism whereby groups are formed endogenously, through the use of voting. Once formed, groups play a public-goods game, where there are economies of scale: in two treatments the social value of an incremental contribution to the group account increases with the size of the group, but in the second treatment, the social value is capped once a certain group size is reached. Societies of nine people are initially formed randomly into three groups of three people who play the game for three periods. Individuals then learn about the average contribution of each individual (by ID number) in one's current own group, as well as the average contribution in other groups, and can decide whether to exit the group. Remaining group members choose whether to exclude any current members from the group; the new groups and 'free agents' then choose whether to merge with other existing groups and/or other free agents.

We find a great degree of success for this mechanism. The average contribution rate is quite high in both treatments, but is modestly (albeit significantly) higher in the first treatment, when there is no cap on the social value of a contribution. In the first treatment, we see large and stable groups forming, but we see considerably more instability and smaller group sizes in the second treatment. The driving force appears to be the economies of scale combined with the awareness that bad behavior will result in ostracism, but in the Athenian sense of possible redemption. This redemption is a unique feature of our environment, with about one-third of the population becoming good citizens after initially being low contributors.

Keywords: Endogenous group formation, Exclusion, Experiment, Merger, Ostracism, Public goods, Social efficiency, Voting

JEL Classifications: C91, A13, B49, C79, C92, D71, H41

Acknowledgements: We thank Nick Burger, David Cooper, Vince Crawford, Martin Dufwenberg, Brit Grosskopf, Andrew Healy, Carlos Oyarzun, Charlie Plott, Andy Schotter, Marie-Claire Villeval, and seminar participants at the 2006 Communication and Incentives Workshop in Santa Barbara, the 2006 North American Regional ESA Meetings in Tucson, and the seminar audience at Texas A&M University for very helpful comments. We thank Nick Burger, Michael Kuhn, and John Lynham for valuable research assistance.

* Charness is at the Department of Economics, University of California Santa Barbara, Santa Barbara, CA 93106-9210, http://www.econ.ucsb.edu/~charness/, email: charness@econ.ucsb.edu, fax: 1-805-893-8830; Yang is at the Research Center for Humanity and Social Sciences, Academia Sinica, Taipei 115, Taiwan, Republic of China, http://www.sinica.edu.tw/~cly/, e-mail: cly@gate.sinica.edu.tw, fax: 886-2-27854160.

1. Introduction

While achieving cooperation is beneficial or even critical for groups or societies, doing so may be problematic since individual incentives often conflict with socially-efficient actions. Even if people feel inclined to help out, they may lose the taste for cooperating when they see other people not doing their share. The issue of how to successfully implement collective action when there is a prospect of such free riding is a vital topic in public economics.

Experiments have investigated the provision of public goods since at least Marwell and Ames (1979). The basic idea is that each person has an endowment to allocate between private goods and public goods, with a contribution to the public good creating more social surplus than a contribution to the private good. However, each individual only receives a fraction of the amount contributed, so that it is rational to allocate one's endowment to private goods. Standard public-goods games typically find moderate initial contributions declining steadily over time.

How can long-term efficiency be sustained? One approach involves identification of individual contributions and voluntary punishment. While some studies (e.g., Fehr and Gächter 2000) show that the level of cooperation can be sustained fairly well, punishment is inefficient, as it involves sacrifice to lower the payoff of another person.[1] Punishment may also undermine altruistic cooperation.[2] An alternative approach is to allow the group in question to evolve endogenously, as with *assortative matching*.[3] Tiebout (1956) argued that efficient provision is feasible if groups were allowed to self-sort, implying a basic mobility to the society.

[1] Masclet, Noussair, Tucker, and Villeval (2003) find that while non-monetary (and so not inefficient) punishments are as effective as monetary punishments in initially increasing contributions, they are less effective over time.
[2] See, for example, Fehr and Rockenbach (2003) and Fehr, Klein, and Schmidt (2007).
[3] Sober and Wilson (1998), Frank (1988), and Bergstrom (2002) show that cooperative behavior can survive if it entails advantages over selfish behavior; these advantages can include an improved chance of being matched with like-minded individuals and a better chance of avoiding undesirable elements of the population. Sethi and Somanathan (2003) provide a survey of theories of assortative matching.

We follow the latter approach in this paper, providing and experimentally testing a more flexible and 'realistic' mechanism for endogenous group-formation that includes possibilities for exit, exclusion, and mergers among groups. The basic structure of our mechanism for group evolution allows for an ebb and flow in a dynamic environment. There are four key features of our design. First, contributions are more valuable in larger groups (up to a point). Second, the possibility of exclusion (or ostracism) affords would-be cooperators insurance that they will be able to ward off would-be free-riders. Third, we permit 'redemption', whereby people who have been initially uncooperative can become good citizens in their society. Finally, a unique feature of our design is that it implements mergers among groups, thus facilitating a more efficient process for combining existing cooperative groups. This allows groups to not only shrink by exit and banishment, but also to grow because their 'cooperative' looks attractive to other groups.[4]

A number of recent papers that have explored the topic of endogenous group-formation in public-goods games, but many or most of our key features are absent from previous work. For example, Erhard and Keser (1999) feature similar efficiency considerations, but without exclusion; the pattern is one of cooperative participants being on the run from free riders, with an overall contribution rate of only 52%. Cinyabuguma, Page, and Putterman (2005) achieve a high rate of contribution among those people not expelled from the original grand coalition. However, exclusion is harsh and irreversible; over time more and more people are banished, with social efficiency eroding. Ahn, Isaac, and Salmon (2006) find that whether or not entry to and exit from groups are unilateral or subject to group consent has substantial effects on behavior and outcome. However, their payoff structure is highly unusual and, in any event, the degree of

[4] In previous studies, groups could only grow by adding individual members. Thus, if two smaller groups wished to combine, one group would first have to dissolve itself.

social efficiency achieved is relatively low: the highest average earnings in any of their treatments were only 42.5% of the social optimum.[5]

A critical element of any group-formation scheme is how one handles the possibility of exclusion. The roots of this notion in human society go back at least to the days of Athens, where a constitutional safeguard called *ostrakismos* is considered to have been an important pillar in the stability of the Athenian democracy. Citizens could vote on whether to banish other citizens who were too powerful or wealthy or otherwise threatened the unity of the society; an individual who received too many votes was banished for a period of 10 years, but could then potentially re-join the society.

Ostracism is evident in almost all civilizations and known cultures (see Gruter and Masters 1986). It may be manifested in informal practices such as 'the cold shoulder' or 'the silent treatment' or be more formal as in Athens (and ancient Syracuse), the Amish practice of shunning, or the American colonial custom described in Nathaniel Hawthorne's *The Scarlet Letter*. Other examples come from the Pathan tribes of Pakistan, the Slavic tribes of Montenegro, and Western Apache culture. Ostracism is also found among animal species such as primates, lions, and wolves (see de Waal 1986, Goodall 1986, and Lancaster 1986). In modern organizations, whistleblowers are often ostracized. A modified and very effective form used with children is a 'time out', where there is (at minimum) a reduction in social attention for a relatively short period of time.

On the one hand, it seems unrealistic to prevent groups from excluding bad actors; on the other hand, a 'grim-trigger' punishment of irreversible banishment seems socially wasteful, and

[5] Some other related studies on this topic include Riedl and Ule (2002), Coricelli, Fehr, and Fellner (2005), Brosig, Margreiter, and Weimann (2005), Maier-Rigaud, Martinsson, and Staffiero (2005), Page, Putterman, and Unel (2005), Croson, Fatás, and Neugebauer (2006), Gunnthorsdöttir, Houser, MaCabe, (2007), and Önes and Putterman (2007).

may also be subject to re-negotiation in a field environment. In general, ostracism describes a custom whereby individuals are excluded in order sustain a group's cohesiveness (see Williams 2001). Typically, the consequences are social and psychological; however, in an anonymous laboratory environment this is less practical, so we instead incorporate financial consequences. However, this exclusion is not necessarily permanent and people can have second chances.

Field examples of mergers between groups are less pervasive. Nevertheless, in the business world, there are joint ventures between research consortiums and mergers between firms, where there are economies of scale involved. In the political realm, there are parallels to environmental advocacy groups, who tend to experience economies of scale, as well as to governing coalitions in parliamentary democracies; we also see cooperation between groups to increase their political influence. In addition, municipal governments may form larger groups for the purpose of sharing the cost of public goods.[6]

Our experiment has two treatments with endogenous group-formation. People play a public-goods game for three periods, during and after which there is at least partial identification (by ID numbers) of the contributions made by others. We then allow exit, ostracism, and merger in three steps, as is described in detail in Section 2. Three periods of play with the new groups, with this process continuing until 15 periods have been played. We have two such 15-period segments, with ID numbers changed for the second segment, thus allowing a fresh start. Each 'society' is comprised of nine individuals, who are initially randomly matched into three groups with three members in each group.

In our first treatment, the total value to the group of the marginal unit of contribution increases with group size, so that larger group sizes are more efficient and the social optimum is

[6] For example, in Lyon, France, there is the so-called "Grand Lyon" that merges together many villages and towns around the city in order to have more extended and diversified network activities in leisure and transportation. We thank Marie-Claire Villeval for providing this example.

achieved with full contributions in the grand coalition of nine people. Our second treatment is identical to the first except that, in order to explore the effect of social efficiency on behavior, we only increase the group's return from contributions up to a group size of four. In both of these treatments, we provide some friction by having the marginal per-capita return (MPCR) of contributing decreases with increasing group size. For our third (baseline) treatment, we also conduct standard public-goods games in which we have fixed groups of size three, six, or nine. The group return (and MPCR) from contributions is the same as for the group sizes in our first treatment, so that we can make comparisons across treatments. This treatment also allows us to rule out that our results in the other treatments are due to some quirk of the UCSB subject pool.

Our results are striking, as we find a great degree of success for this mechanism in both of our treatments where it is used. We see considerable exit and exclusion, particularly in early rounds. The average contribution rate (excluding the final three periods of a segment) is quite high in our first two treatments. When rewards from increasing group size are not truncated, large and stable groups form, promoting efficiency: With truncation, contributions are slightly lower, with smaller groups and considerably more instability. The high level of contribution persists while subsequent voting for group members is expected. However, we do find a significant end-game effect, as contributions in the last three periods of the second segment (after the last voting round) drop dramatically; this suggests that it is not simply a taste for altruism that drives our results. Our baseline treatment shows nothing unusual in our subject pool, with much lower levels of contribution and the familiar pattern of a steady decline over time.

The remainder of this paper is organized as follows. In Section 2, we describe the experimental implementation and provide hypotheses; we present the experimental results in Section 3. Section 4 offers a discussion of our results and we conclude in Section 5.

2. Experimental Implementation and Hypotheses

2.1 The Experiments

We conducted our experiments at the University of California at Santa Barbara, with participants recruited from a campus-wide list of (primarily) students who had previously indicated interest in participating in paid research experiments. None of the participants had any experience with public-goods or voting experiments.

We used a control treatment (Treatment 1), which featured fixed groups, in order to establish that our participants show similar behavior in the standard game to what has been observed previously. There were three sessions with 18 people in each, so 54 participants in all. People were randomly assigned to a 3-person group, a 6-person group, or a 9-person group for a 15-period segment. We chose these size groups to get observations from small, medium and large groups. Each group simultaneously played the same public-goods game as in our main treatment, except that there was no endogenous group formation and no information was provided about the contributions of the other individuals or groups.[7] After the first segment, we randomly re-assigned people to new groups, so that the composition of groups varied across segments. Average income in these 45-minute sessions was about $13.

For our main treatment (Treatment 2), we conducted four computerized laboratory sessions with a web-based program; there were nine participants in the first session and 18 in each of the remaining three sessions, for a total of 63 participants in the endogenous group-formation sessions. The average income in these 90-minute sessions was approximately $21, including a $5 show-up fee. Our supplemental, printed instructions (for Treatment 2) are

[7] We are indebted to Charles Holt for adapting his Veconlab software to make this possible.

presented in Appendix A.

Nine people in each 'society' were placed into three groups of three people and played a public-goods game for three periods. Each person was endowed with 25 tokens that could be allocated between a private account and a public account. The social value of an allocation to the public account depended on the group size, as shown in Table 1. We determined the group return by using the four-person MPCR of 0.4 as the standard, leading to a group return of 1.6 for the 4-person group. We increased (decreased) this factor by 10% for each additional (fewer) person in the group, except that the factor for a solitary individual was 1.

Table 1: Group returns from contributions and MPCR in Treatment 2, by group size

n	1	2	3	4	5	6	7	8	9
Group return	1	1.322	1.455	1.600	1.760	1.936	2.130	2.343	2.577
MPCR	1	0.661	0.485	0.400	0.352	0.323	0.304	0.293	0.286

After the first three periods, individuals learned about the average contribution of each individual (by identification number) in one's current group in those three periods, as well as the average contribution in other groups. At this point, we proceeded with exit and exclusion. First, people could unilaterally decide whether to exit the group. Each person who elected to stay in their group voted on every other group member; if more than 50% of the voters chose to exclude an individual, that individual was expelled from the group.

At this point there were a number of groups, group remnants, and 'free agents' (individuals who were unmatched either voluntarily or not). Each of these entities then decided whether it wished to merge with each of the other entities, with information provided about the average contribution for the last three periods for each entity.[8] A merger happened if and only if at least 60% of the members of each entity wish to merge with the other entity.[9]

Our algorithm for performing group mergers began by considering whether the two largest groups wished to merge. If not, the mechanism considered the desires of the largest group and the third-largest group regarding a merger.[10] This continued until a merger was achieved, at which point the process began again with the new two largest groups, etc. Ultimately, the process continued until no more mergers could be made. If there were leftover singles, these were formed into new groups, depending on how many singles remained.[11]

The new groups then played three periods of the public-goods game with the same rules. There were voting rounds after the first four sets of three periods, with 15 periods overall in the segment. We then re-formed societies (except in the first session) with new identification numbers randomly assigned; a new 15-period segment was started with three three-person groups formed at random in each society.[12] Thus, in each session there were 30 periods of play in the public-goods game and eight voting rounds.

We note that since we have a pre-determined number of periods, the only subgame-perfect equilibrium is non-contribution. To see this, note that everyone with selfish preferences

[8] This also implies that subjects may intentionally display their previous contribution level by choosing to exit. In a field environment, the natural data observable are the personal incomes they received from their previous firms or associations, in the spirit of Tiebout. However, given the previous group's level of public goods and the single's income, her average contribution can easily be calculated, and thus is informationally equivalent to her income.

[9] This sequence seemed the most natural to us. Of course, other sequences would also have been feasible, such as instead allowing exit after the mergers have been determined, in case a minority voter was unhappy with a merger. However, this would limited the supply of free agents at the time of the mergers, and exit at this stage would have been quite costly for people choosing to leave (since they couldn't join any existing groups of cooperators).

[10] All ties in size ranking were broken randomly

[11] If there was exactly one single at the end of this process, he or she remained unattached. The details of this process can be found in the Instructions provided in Appendix A.

[12] We chose to have a re-start, as an individual might get locked into a situation that is difficult to escape during a segment, but instead would receive a fresh start in the second segment. We allowed the societies from one segment to mix in the second segment for at least two reasons. First, by comparing the patterns of behavior across the two segments, we can assess how much was learnt in the first segment. Second, mixing the groups across segments also helped to avoid the resumption of the behavioral patterns from the first 15-period segment.

will contribute nothing in the last three periods of a segment. Anticipating this, group size has no value in these periods, so everyone would contribute nothing in the penultimate three periods, etc. In other words, in the standard game-theory framework, our setting is essentially not different from the one without our mechanism.[13] Of course, if we convert our game to one with a stochastic ending, we could sustain cooperation.

After having observed the results of our initial treatment, we designed an additional treatment (Treatment 3), to isolate the influence of efficiency considerations on contributions and group size. This treatment was identical to our initial treatment, with one important exception: the group return from contributions was set at 1.6 for all groups with four members or more. We had three sessions in this treatment, each with 18 participants. The average income in these 90-minute sessions was approximately $19, including a $5 show-up fee. Table 2 shows the value of an allocation to the public account; it is clear that in this case there is no direct efficiency incentive to form larger groups.

Table 2: Group returns from contributions and MPCR in Treatment 3, by group size

n	1	2	3	4	5	6	7	8	9
Group return	1	1.322	1.455	1.600	1.600	1.600	1.600	1.600	1.600
MPCR	1	0.661	0.485	0.400	0.320	0.267	0.229	0.200	0.178

2.2 Hypotheses

We present six behavioral hypotheses in relation to the patterns in our data and comparisons across treatments. As we naturally presume that our mechanism will have some degree of effectiveness, our first hypothesis is simply:

[13] This is essentially true for all VCM experiments with added features ranging from punishment to ostracism to endogenous sorting.

H1: There will be a significantly higher contribution rate in both of our endogenous group-formation treatments than in our control treatment.

We also expect a substantial amount of exit, exclusion, and mergers; further, these choices will be correlated with relationships between the contributions of the various parties:

H2: One will be more likely to choose exit the larger is the difference between one's own contribution rate and the contribution rate for one's group, and the larger is the average contribution of other groups.

H3: One will be more likely to exclude a group member the lower is that group member's contribution rate.

H4: One will be more likely to vote for one's group to merge with another group the larger is the other group's size and average contribution.

Finally, we expect that there will be a strong role for efficiency (the group return). This leads us to two predictions across our treatments with endogenous group-formation. Since the potential increase in efficiency is much greater in Treatment 2, and since there is no efficiency gained by having groups larger than size four in Treatment 3, we have:

H5: Average contribution rates will be higher in Treatment 2 than in Treatment 3.

H6: There will be more groups with at least five members in Treatment 2 than in Treatment 3.

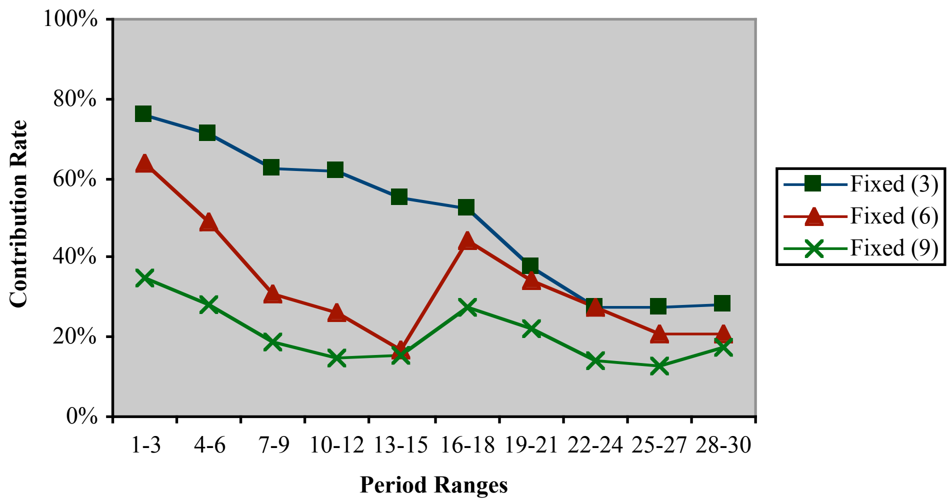
3. Experimental Results

In this section, we first present descriptive statistics and charts about our results. We then proceed to test our hypotheses using nonparametric statistics and regressions, and conclude the section with a discussion of efficiency and profits.

3.1 Descriptive statistics and charts

Figure 1 illustrates the average contribution rate over time with fixed groups:

Figure 1: Average contribution rate over time, Treatment 1

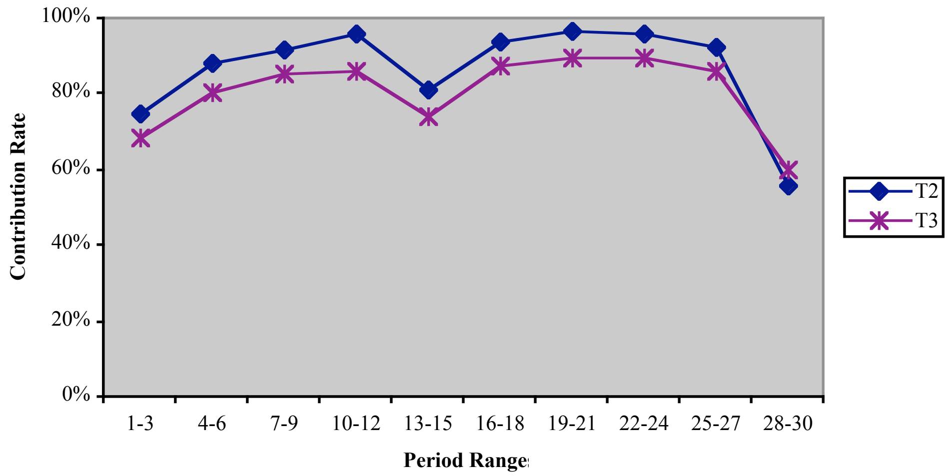


We see the familiar pattern of a moderate initial rate of contribution, followed by a decline over time; by the end, the overall rate of cooperation is only about 20-30%. Note that there is a consistent effect of the MPCR across the fixed-group treatments, particularly in the first 15-period segment. There is a substantially higher contribution rate in the 3-person game (MPCR = 0.485) than in the 6-person game (MPCR = 0.323), and a substantially higher contribution rate in the 6-person game than in the 9-person game (MPCR = 0.286). The order of these rankings is preserved in the second segment, but the differences are much smaller. We do observe a major re-start effect in the 6-person game, with a smaller one in the 9-person game.

Figure 2 shows the average contribution rate over time for Treatments 2 and 3, excluding unmatched players (singles) from the calculations, as their contribution choices are arbitrary.[14]

[14] There were only 24 cases in which an unmatched player chose a contribution out of 1170 cases in total, or about 2%. Thus, even assuming a zero-contribution rate for these singles would not affect the results substantially.

Figure 2: Average contribution rate over time, Treatments 2 & 3



The average contribution profiles over time are remarkably similar for Treatments 2 and 3; this behavioral feature surprised us. Our mechanism induces high contribution levels, with a notable decrease in the last voting period in a segment.[15] There is a 5-10 percentage point reduction in contributions in every period range except the last one; overall, the contribution rate in Treatment 2 (excluding singles and the last voting period) is 90.9% (87.2% in segment 1 and 94.5% in segment 2), while the contribution rate in Treatment 3 is 83.9% (79.8% in segment 1 and 88.0% in segment 2).

Clearly endogenous group formation enhances contributions to the public good in comparison to fixed groups. In the first set of periods of Treatments 2 and 3, where the group size is always fixed at three, the contribution rate (74% in Treatment 2 and 68% in Treatment 3) is nearly identical to the contribution rate with fixed three-person groups. However, this rate shrinks steadily to around 25% with fixed groups, while it grows to near or above 90% in all sets of periods in Treatment 2 and at or above 85% in Treatment 3 (except for the last voting period

[15] We use "round" and "voting period" interchangeably. Each refers to the three periods of choosing contributions, as well as the three voting stages after these (where this is applicable in the context).

in the segment). Note the dramatic end-game effect in both Treatments 1 and 2 in periods 28-30, indicating that good behavior doesn't appear to be intrinsic; this drop is also echoed in periods 13-15.[16] In theory, we should observe unraveling all the way back to the beginning, but it is more the norm than the exception in experiments to see unraveling only in the last period or two of finite-horizon experiments.[17]

Table 3 considers the average contribution by group size, for both the endogenous and fixed-group treatments. We omit the last three periods of each segment in the endogenous treatment, since these take place after the voting has ended, and also do so for the fixed-group treatments, for parallelism.

Table 3: Average contribution, by group size

	Group Size							
Treatment	2	3	4	5	6	7	8	9
1	-	13.02 (72)	-	-	9.29 (144)	-	-	5.42 (216)
2	19.38 (56)	20.21 (150)	25.00 (4)	24.55 (40)	25.00 (42)	24.17 (49)	24.67 (80)	24.91 (72)
3	19.76 (80)	19.61 (180)	23.77 (84)	22.23 (40)	22.70 (18)	19.71 (7)	23.40 (16)	-

The number of observations is in parentheses.

In Treatment 2, we see that the average contribution is very high (at least 96.7%) for all group sizes of four or more. The rate is appreciably lower for two-person and three-person groups (although still greater than 77.5% in every case for both endogenous groups); the reason for this is that the two-person groups are often formed from unmatched singles after the merger stage, and three-person groups are mainly observed in the first three periods of a segment, where

[16] Perhaps in these periods, as opposed to periods 28-30, some people were concerned about whether their behavior would somehow be identifiable in the subsequent periods, and so didn't want to risk having their selfish behavior in the first segment exposed in the second 15-period segment. Alternatively, people might have been learning an evolving social norm. After the first segment, they learned to begin with a high contribution level and to form large groups right away, but also realized that free-riding in the last periods had no adverse consequences.

[17] We do observe some people reduce their contributions heavily in the penultimate round of a session.

(at least in the first segment) contributions are lower. Nevertheless, Treatment 3 does not feature such high contribution rates (although this rate is 95% for 4-person groups).

Since a primary feature of our design allows group size to be endogenous, we consider these patterns, as well as how each 9-person society evolves over time. Table 4 shows the frequency of the various components of the societies in Treatment 2 over time:

Table 4: Frequency of decompositions of the 9-person societies in Treatment 2 over time

	Decompositions								
	(2,2,2,3)	(3,3,3)	(2,3,4)	(2,2,5)	(1,2,6)	(3,6)	(2,7)	(1,8)	(9)
Periods 1-3		7							
Periods 4-6	1		1	4		1			
Periods 7-9				3	1		3		
Periods 10-12				1			2	4	
Periods 13-15					1		2	1	3
Periods 16-18		7							
Periods 19-21						2	1	2	2
Periods 22-24						2		2	3
Periods 25-27						1	1	2	3
Periods 28-30						1	2	2	2
Total	*1*	*14*	*1*	*8*	*2*	*7*	*11*	*13*	*13*

It is clear that group size grows steadily over the course of the first 15-period segment; for example, while there were no groups of seven or larger in periods 4-6, there were three in periods 7-9, and six in both periods 10-12 and periods 13-15 (seven is the maximum possible, one for each of the seven societies). It is also apparent that group sizes are larger in the segment after the re-start, as there were five groups of seven or larger in both periods 19-21 and periods 22-24, and six in both periods 25-27 and 28-30. Overall, the most common endogenous decompositions were in fact the grand coalition and (1,8), followed closely by (2,7); in the second segment the grand coalition was the most common endogenous group.

Table 4 also confirms the strong move towards large groups as the segment continues,[18] and it also reveals something about the evolution of cooperation via learning/education. The (2,3,4) and (2,2,5) groups in periods 4-6 show that some singles exited their initial groups and were then accepted into a more cooperative group, likely via revelation of their relatively high contribution level when they became free agents. This helped to create building blocks for further movement towards larger groups and consistently high levels of contribution, as people wish to signal willingness to conform to the norm in order to be accepted into the better 'club'.[19]

Table 5 shows the decomposition patterns in Treatment 3:

Table 5: Frequency of decompositions of the 9-person societies in Treatment 3 over time

	Decompositions										
	(2,2,2,3)	(3,3,3)	(1,2,2,4)	(2,3,4)	(1,4,4)	(2,2,5)	(1,3,5)	(1,2,6)	(3,6)	(2,7)	(1,8)
1-3		6									
4-6				4		1		1			
7-9				3	1	1	1				
10-12				3		1			1		1
13-15	2		1	2				1			
16-18		6									
19-21		1		3		1					1
22-24	1			1	1	2		1			
25-27	1			3		1				1	
28-30	1	1		2		2					
Total	*5*	*14*	*1*	*21*	*2*	*9*	*1*	*3*	*1*	*1*	*2*

The pattern differs considerably from the pattern in Table 4. By far the most common decomposition (especially leaving aside the mandatory start in the first round) is (2,3,4); (2,2,5) is a distant second. As there is no direct incentive to form a group with more than three other

[18] Overall, we see that in 54 of 56 cases (96%), excluding the first three periods of a segment (where 3-person groups were mandated), at least a 5-person group was formed. The grand coalition was formed in 10 of 28 cases (36%) in the second segment.

[19] In a sense, our design has the flavor of a tournament; we thank David Cooper and Andrew Schotter for pointing this out. At first, it is because you (or your group) are trying to get picked in the beginning. Later, it is more like having growth opportunities if you are deemed a reasonable partner in the joint venture.

people (there may be an ‘insurance’ motivation for being in a group of five, if one person starts behaving badly), in Treatment 3 we observe no real trend towards increasing group size. The frequency of various group sizes is revealing: There were 26 groups of size four in Treatment 3, compared to only one in Treatment 2; there were 10 groups of size five in Treatment 3, compared to eight in Treatment 2. However, while there were 46 groups of size six or greater in Treatment 2, there were only seven such groups in Treatment 3.

Another relevant issue is the stability of the groups across treatments. Table 6 below shows the number of larger groups (with four or more people in Treatments 2 and 3), how many of these groups survived intact into the next round, and the number of exits and exclusions in the first four voting periods of each segment.

Table 6: Stability of larger groups

		Group size						
		4	5	6	7	8	9	All
Treatment 2	Total	1	8	7	7	10	8	41
	Intact	1	7	7	4	7	6	32
	Exits	0	0	0	1	1	0	2
	Exclusions	0	1	0	2	2	2	7
Treatment 3	Total	21	8	3	1	2	0	35
	Intact	12	3	0	0	0	0	15
	Exits	4	2	2	2	2	0	12
	Exclusions	7	8	2	1	2	0	20

“Intact” means that the group no person left the group. We count exclusions even when exit was chosen.

Overall, 78% of the larger groups in Treatment 2 stayed intact, while only 43% stayed intact in Treatment 3.[20] Groups of size four were somewhat stable in Treatment 3, but the survival rate for these groups was still only 57%; groups of size five were sometimes stable, but groups larger

[20] Four of the seven break-ups in Treatment 2 occurred at the end of the fourth voting period of the segment, as some people started reducing contributions in the fourth round, perhaps anticipating a collapse in the final round.

than this were *always* unstable in Treatment 3. In contrast, groups with more than five members were stable 75% of the time in Treatment 2.[21]

Comparisons among smaller groups are more problematic, particularly since most of these are formed with unwanted free agents (or are in the initial set-up). While most small groups are unstable, the break-up rate depends heavily on whether or not the group was formed voluntarily. With forced matching, we observe instability for 68% and 77% of the two-person groups in Treatments 2 and 3, and for 100% and 92% of the three-person groups in Treatments 2 and 3. This compares to break-up rates of 33% and 39% for voluntary two-person groups and 0% and 42% for voluntary three-person groups in Treatments 2 and 3, respectively.

As mentioned earlier, we find that many people achieve redemption, as evinced by a pattern of small initial contributions and larger contributions in late rounds. These are people who presumably would have been banished permanently in studies featuring irreversible exclusion. We present each individual's contributions over time and classification with respect to redemption in Appendix B (Tables B1 and B2). While the classification is necessarily somewhat arbitrary,[22] the pattern is nevertheless clear. Overall, 39 of the 117 participants were rehabilitated, with 19 (of 63) of these in Treatment 2 and the other 20 (of 54) in Treatment 3. These proportions are not significantly different; the test-statistic is $Z = 0.79$.

3.2 Hypotheses tests

Based on a comparison between Figure 1 and Figure 2, it is apparent that the contribution rate is much higher with endogenous group-formation than with fixed groups, regardless of how

---

[21] It turns out that of the 14 exits from larger groups, eight occurred in the penultimate voting period, again perhaps in anticipation of unraveling in the final round of the segment. It also turns out (for all group sizes) that people were only excluded when their contribution was less than the average contribution in the group.

[22] To qualify for our definition of redemption, a subject's average contribution must be below 15 in either the first round or the first two rounds of the first segment, and his or her contribution is no less than 21.3 (or more usually, greater than 24) in the third or fourth rounds of the second segment.

the comparison is made. Regarding H1, a simple and conservative statistical test uses session-level data. The average contribution in each of the four separate sessions of Treatment 2 and each of the three sessions in Treatment 3 is (far) higher than the average contribution for any of the three groups in each of the three different fixed-size groups. Making pairwise comparisons between each endogenous treatment and each of the fixed-group-size treatments, the Wilcoxon-Mann-Whitney ranksum test (see Siegel and Castellan 1988) indicates statistical significance at $p = 0.029$ (with the indicated one-tailed test) in each case for Treatment 2 versus Treatment 1 and $p = 0.050$ (one-tailed test) in each case for Treatment 3 versus Treatment 1.[23]

To test for exit, exclusion, and merger patterns (H2-H4), we use random-effect probit regressions. Table 7 shows the determinants of exit choice, excluding exit decisions by unmatched singles:

Table 7 – Determinants of Exit Choice

Independent variables	(1) Treat 2	(2) Treat 2	(3) Treat 2	(4) Treat 3	(5) Treat 3	(6) Treat 3
Own contribution – average ingroup contribution	0.176*** [0.044]	0.152*** [0.045]	0.104** [0.044]	0.071*** [0.025]	0.099*** [0.030]	0.076** [0.031]
Average contribution of other groups		0.094*** [0.033]	0.121** [0.054]		0.101*** [0.031]	0.077** [0.033]
Maximum contribution of other groups			0.207*** [0.032]			0.156*** [0.026]
Constant	-2.034*** [0.213]	-3.920*** [0.787]	-5.103*** [1.321]	-1.158*** [0.130]	-3.237*** [0.663]	-3.304*** [0.727]
# Observations	493	493	493	425	425	425
Log-likelihood	-122.04	-113.03	-79.20	-180.79	-171.56	-148.82

Standard errors are in brackets. **, *** indicate significance at the 5% and 1% level, respectively

[23] While it is true that we don't provide information concerning the contribution of each individual group member in the public-goods game version in Treatment 1, the very large differences in contribution rates across the endogenous and fixed-group treatments and the fact Croson (2001) finds no difference in average contributions according to whether or not this information is provided causes us to strongly suspect that our test result would hold in any case.

In both treatments and in all specifications, the difference between one's own contribution and the average ingroup contribution is a significant factor: the bigger this difference, the more likely that exit will be chosen. When contributions in one's own group are low, one might be tempted to seek more fertile ground. In this respect, the average contribution for the other existing groups is also a statistically-significant determinant of the exit decision. Thus, our results provide strong support for H2. In addition, there is an attraction to the maximum average contribution of the other groups, as we see a large coefficient for this term in specifications (3) and (6), as well as a high degree of significance. A person who chooses exit aspires to be in the best group around.

Turning next to exclusion patterns (H3), if a person has chosen to stay with the group, but is less than thrilled with the choices made by some of the group members, another choice to consider is exclusion. Table 8 reports the results of random-effects regressions vis-à-vis the decision of whether or not to exclude another group member:

Table 8 – Determinants of Exclusion Choice

Independent variables	(1) Treat 2	(2) Treat 2	(1) Treat 3	(2) Treat 3
Contribution	-0.205*** [0.015]	-0.116*** [0.017]	-0.191*** [0.015]	-0.143*** [0.016]
Contribution less than own		1.652*** [0.170]		1.185*** [0.139]
Constant	2.919*** [0.329]	0.421 [0.414]	3.127*** [0.316]	1.542*** [0.356]
# Observations	2089	2089	958	958
Log-likelihood	-315.82	-266.00	-344.25	-302.42

Standard errors are in brackets. *** indicates significance at the 1% level.

We see quite strongly that an individual is less likely to exclude another group member the higher that member's contribution; this is the case in both treatments and all specifications, providing support for H3. In addition, one's own contribution plays a role – if the other person's contribution is less one's own, exclusion is considerably more likely.

Once all exit and exclusion choices have been made, there is the choice of whether to merge with other groups. Table 9 shows the determinants of the individual merger decision:

Table 9 – Determinants of Merger Choice

	Dependent Variable			
Independent variables	(1) Treat 2	(2) Treat 2	(3) Treat 3	(4) Treat 3
Outgroup contribution	0.127***	0.139***	0.153***	0.188***
	[0.010]	[0.010]	[0.010]	[0.011]
Outgroup size	0.245***	0.217***	0.236***	0.137***
	[0.041]	[0.043]	[0.040]	[0.042]
Own group contribution		-0.037***		-0.073***
		[0.013]		[0.010]
Own group size		-0.061**		-0.030
		[0.030]		[0.036]
Constant	-2.965***	-2.174***	-3.748***	-2.131***
	[0.192]	[0.274]	[0.204]	[0.257]
# Observations	1147	1147	1830	1830
Log-likelihood	-560.13	-548.58	-884.49	-787.19

Standard errors are in brackets. **, *** indicate significance at the 5% and 1% levels

A merger is more attractive if the other group has made higher contributions; further, the larger the other group, the more interest there is merging with it. These forces are quite significant and strong, providing support for H4. In addition, the smaller the average contribution in one's own group and the smaller one's own group size, the more attractive a merger becomes for the members of the group.

We turn now to the hypotheses concerning comparisons between Treatment 2 and Treatment 3; what is the effect of truncating potential efficiency gains at a group size of four? The results are clear concerning the number of large groups in each of these treatments (H5).

Referring back to Table 4, there were 46 instances in Treatment 2 where there was group with five or more members, out of the 56 possible cases (so 82% of the time). In contrast, Table 5 shows that there were 17 instances in Treatment 2 where there was group with five or more members, out of the 48 possible cases (so 35% of the time). Session-level data provide support for H5, as the number of large groups was higher in each of the four sessions of Treatment 2 than in any of the three sessions of Treatment 3. The Wilcoxon-Mann-Whitney ranksum test indicates statistical significance at $p = 0.029$ (with the indicated one-tailed test). Thus, larger groups do not tend to form when there is no payoff advantage for the larger groups.

The results for our hypothesis concerning the effect of efficiency truncation on contribution rates (H6) are more mixed. Recall the strong similarity in the contribution profiles over time in Figure 2, with a rather modest difference in contribution rates in every period range. Table 10 shows the average contribution rate in each society, excluding singles and the last voting period of each segment:

Table 10: Average contribution by society and treatment

	Session 1		Session 2		Session 3		Session 4		
Group	S1	S2	S1	S2	S1	S2	S1	S2	Overall
T2, 1st seg.	20.68	-	21.82	22.90	21.90	19.56	22.27	23.57	21.80
	Session 5		Session 6		Session 7				
T3, 1st seg.	17.42	22.27	21.85	17.98	18.67	21.82			19.96

	Session 1		Session 2		Session 3		Session 4		
Group	S1	S2	S1	S2	S1	S2	S1	S2	Overall
T2, 2nd seg.	24.84	-	24.95	24.17	23.47	18.48	25.00	24.67	23.64
	Session 5		Session 6		Session 7				
T3, 2nd seg.	22.89	23.24	21.09	20.50	21.19	23.08			22.00

Once again, the ranksum test can be performed using each society as one observation.[24] The purest test uses data from the first segment, as there has been no interaction between societies. However, the data from the second segment has the advantage of being more 'mature', perhaps better representing settled behavior. The Wilcoxon-Mann-Whitney test yields $p = 0.069$ (one-tailed test) using society averages in Segment 1, and it yields $p = 0.018$ (one-tailed test) using society averages in Segment 2. A similar approach that instead uses the average contribution of the median contributor in each society as one observation, also finds significance: $p = 0.051$ (0.001) for one-tailed tests of the data from segment 1 (2). So there is at least some support for H6; nevertheless, the difference in rates across treatments is quite modest and is certainly smaller than we expected. The mechanism achieves a high level of contributions even when the potential efficiency gain is much smaller (0.60 versus 1.57).

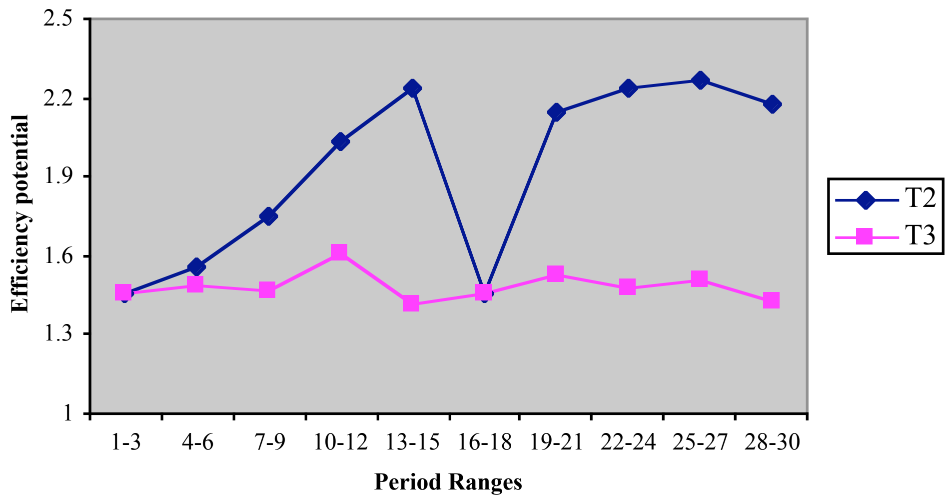
3.3 Efficiency and profits

How well do our societies do in terms of achieving efficiency? Of course, this is related to the contribution rate, but another factor in our design is the group size. A useful metric is the weighted efficiency potential *(EP)*. A nine-person society with a group of 8 and a group of 1 has higher potential efficiency than the same society comprised of a group of 5 and a group of 4. This variable implicitly summarizes the distribution of group sizes in the society. We derive this measure as follows: Let the society consist of groups 1, 2, …, N with sizes, s_1, s_2, ... s_N. Let f(k) denote the multiplying factor of a group of size k. Then, the $EP = [f(s_1)*s_1+ f(s_2)*s_2 + ...)]/(s_1+$

[24] We also perform random-effects regressions with the contribution as the dependent variable and the treatment (2 or 3) as an independent dummy variable. With clustering on the individual, the coefficient on the treatment dummy has a test-statistic of $Z = 2.63$, $p = 0.004$ (one-tailed test); with clustering instead on the society in a segment, the coefficient on the treatment dummy has a test-statistic of $Z = 1.95$, $p = 0.026$ (one-tailed test).

$s_2+...) = [f(s_1)*s_1+ f(s_2)*s_2 + ...)]/9$ in our design.[25] Figure 4 shows the evolution of the efficiency potential over time:

Figure 4: Efficiency potential over time



In periods 1-3 and 16-18, the *EP* is necessarily 1.455, as we start with three groups of three in each segment. In Treatment 2, the *EP* grows steadily through the first segment, reaching over 2.2 by the end of the segment; the growth is fairly immediate in the second segment, as the *EP* is already 2.15 in periods 19-21; it stays in this vicinity during the remainder of the second segment. Treatment 3 shows a very different pattern, as here the efficiency potential (which is capped at 1.60) is largely constant over the course of the sessions.

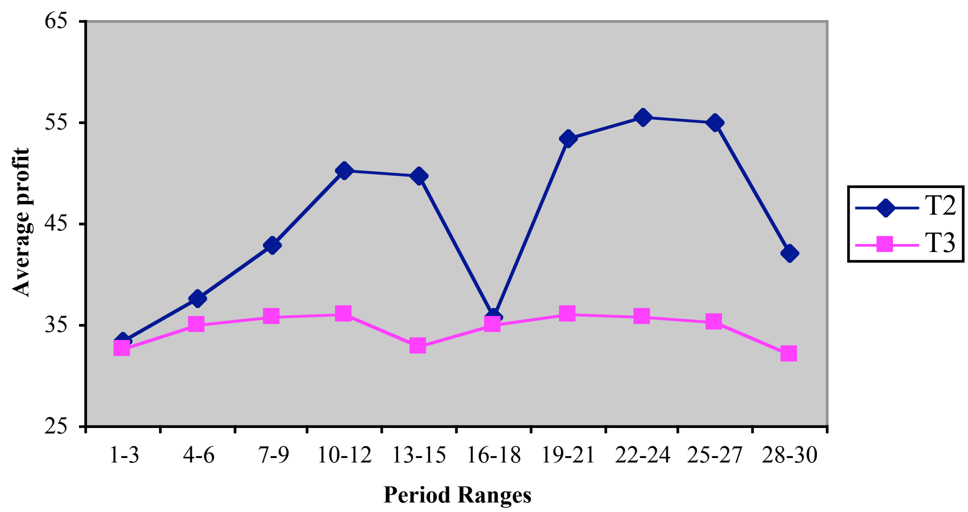
How does this increase in group size and efficiency potential come about in Treatment 2? First, we observe substantial amounts of both exit and exclusion. A participant typically exits a group when his or her contribution exceeds the group average; it is the quickest way to sever ties with a group of potential free riders. Exclusion tends to take place in the first voting period,

[25] Given the numbers in Table 1, the highest possible *EP* in Treatment 2 is 2.577, for the grand coalition. If there are three groups of three (the initial set-up), the *EP* is 1.455. The most efficient formation in Treatment 3 is either a combination of a group of four and a group of five, or the grand coalition. In either case, the *EP* is 1.6. Note that we never observe the most efficient formation in Treatment 3, but see it 13 times in Treatment 2.

when one of the three group members has made relatively low contributions, or in later voting periods, when some people in a large group make small (or no) contributions. But the major force in expanding group size is the innovation of our merger process between groups. There were 23 mergers between groups in Treatment 2, with at least one in all but one of the 14 societies in Treatment 2; 16 of these mergers occurred at the end of the first round, getting the growth process jump-started (and 11 of these were in the second segment, suggesting subjects had learned the value of merging early). By contrast, there were only 11 such mergers in Treatment 3, with six of these occurring at the end of the first round (three in each segment; perhaps subjects had not learned of any great value in merging early). Table C1 in Appendix C shows group sizes and mergers over time for each society in each treatment.

To translate the efficiency potential into overall profits, we combine it with the average contribution made. Figure 5 presents the realized profits that are achieved over time in our endogenous sessions.

Figure 5: Realized profits over time



In Treatment 2, we see that profits tend to increase over time, except during the last three periods of a segment; note that the average realized profit in the last three periods of the second segment of Treatment 2, even with the considerable free-riding, is much higher than that obtained in the first three periods of either segment in that treatment. In fact, to achieve the same level of average profit in the first three periods of a segment as in these last three periods of the second segment, the people in the 3-person groups would have to contribute an average of 97% of their endowments. In Treatment 3, profits don't change much over time, averaging around 35 (40 was the maximum possible average), but with small dips in the last round of both segments.

Our final question in this section is the relationship between one's own contribution rate and one's profits: Does cooperation pay? This is the basic evolutionary question (and assortative-matching issue) to be answered. Yang, Yue, and Yu (2007) show there is a positive correlation between average payoff and how cooperative a player is overall. Of course, holding one's group size and the contributions of others in the group constant, the higher the contribution, the lower the profit. But it is obvious that contributions are a critical determinant of the size and character of one's (endogenous) group. In Treatment 2 (excluding unmatched players), we obtain a Spearman coefficient of 0.3952, with $p = 0.000$; in Treatment 3, the Spearman coefficient is 0.2106, again with $p = 0.000$. In Appendix D, we show the visual relationship between average contributions and profits for each individual in Treatments 2 and 3.

4. Discussion

Our mechanism is successful in achieving a high rate of contribution in both treatments. In Treatment 2, we find a pronounced tendency towards endogenously *forming* large, full-contribution groups, thus promoting a high degree of efficiency. We often see the grand coalition formed, even though we start with relatively small building blocks and have only a few voting rounds to get there. This is clearly due to the possibilities for exit, exclusion, and mergers, since cooperation dissipates when these actions are no longer possible, both in the last three periods of a segment and in our control experiments. In Treatment 3, contributions are slightly (but significantly) lower when we truncate (with respect to group size) the social value of an individual contribution, resulting in considerably less stability and overall efficiency.

Our results are driven by three major elements. First, to form large and inclusive groups, it is necessary to provide returns to scale; if these increasing returns to scale cease above some group size, the rare groups larger than this size are rather unstable. The presence of economies of scale seems to be a natural feature for public-goods problems in general. Hence, it is important to deal with truly endogenous group sizes. Second, it is critical that would-be cooperators understand that they have insurance against 'selfish' and myopic behavior by people in their group. Third, people often learn (and recover) from early mistakes in judgment, and become rehabilitated or redeemed; this feature is closer to the Athenian *ostrakismos*.[26] The underlying society is of course more effective and cohesive when almost everyone pitches in. Redemption is an indirect byproduct of a successful assortative-matching institution.

While redemption looks at revealed changes in individual behavior, a strong (and rationalizable) belief that a grand coalition or other large group will form may suppress one's inclination towards free-riding right from the start. This seems much more the case in T2 than in T3, as there are still many subjects in T3 with average contributions below 20 over the entire 30 periods, but there are rarely such cases in T2; this is evident from Figures 6 and 7. In comparison with Page, Putterman, and Unel (2005), for example, the preponderance of the one-

[26] Ohtsuki and Iwasa (2006) found that "apology and forgiveness" is a necessary property for any sound cooperation-promoting mechanism in the image-scoring context.

quarter of the population revealed as 'free-riders' and placed in their lowest-contributing group seem to be converted into 'model citizens' in our endogenous treatments.

Unlike previous studies involving exclusion, we achieve our high rate of cooperation and efficiency without the very costly 'grim-trigger' punishment of permanent expulsion, as an excluded person is often rehabilitated (redeemed), becoming a highly-contributing member in another group or even the original group.[27] Also, while cooperation is more valuable in larger groups, there is at least some congestion, so that the MPCR does drop with increasing group size, unlike the other studies with exclusion. Thus, in a certain sense our design represents a more difficult test for sustained cooperation, as the burden increases with group size.

Another key issue concerns the degree of information required to achieve successful outcomes. Nearly all of the other studies mentioned provide complete information regarding the contributions (or effort) provided by each of the participants in the session. This is a strong requirement, which may not be realistic in the field environment. We feel that it is more in the spirit of Tiebout (1956) to have information shielded by the boundaries of the groups; while one might be able to observe the total output of a group or firm to which one does not belong (for example, from public accounting or tax records), it is more difficult to justify an outsider being able to learn the detailed production of each person in a group or firm. In our environment, we do not allow the inner workings of each group to be exposed as long as the group is intact.

How can we explain the observed behavior? After all, in theory we should observe unraveling all the way back to the initial period, with zero contributions throughout; this is the same prediction for our fixed-group treatment. Yet a robust phenomenon in standard public-goods experiments is to have a substantial proportion of contributions in the beginning, followed

[27] One might argue that permanent exclusion is a 'grim-trigger' strategy that is perhaps not the best match for the field environment, where people tend to eventually relent or to renegotiate.

by considerable decay over time; this is also what we observe in our Treatment 1. Any behavioral story applying to our endogenous-group treatments must also be relevant for the fixed-group treatment.

The main issue in the standard public-goods game involves efficiency versus selfish behavior. Charness and Rabin (2002) show that people often have a taste for social efficiency, defined as the total group payoff. Thus, many people will be willing to make large contributions, at least initially. However, to the extent that cooperators notice that others are not contributing (either individually or in the aggregate), they withdraw their concern for the misbehaving other people (or the aggregate group). Depending upon one's weight on others' payoffs, sufficient misbehavior will make sacrifice less attractive and a threshold (varying across individuals) may be reached where one refuses to contribute.[28]

This fits well with the explanation offered by Fischbacher and Gächter (2006), which is based on the presence of different types in the population. First, there are some unconditional cooperators who contribute all or almost all of their endowment, regardless of the circumstances. Second, there is some proportion of *conditional* cooperators, who cooperate as long as others are doing so.[29] Third, there are free-riders, who make small or no contributions from the outset; it is the presence of these free-riders that drives the decay in contributions over time.

But the dynamics are considerably different in our endogenous group-formation treatments. The key is that free-riders are not able to exploit cooperators on a long-term basis, as these free-riders are excluded from cooperative groups until they reform. Conditional cooperators will try to avoid free-riders and, if successful, will make large contributions; they

[28] Andreoni and Samuelson (2006) make a similar point, in twice-repeated prisoner-dilemma games in which "relationship design" (weight on first versus second period) multiplies the beneficial effects of altruism and reciprocity.

[29] Our results in the final round suggest that some people may be conditional cooperators until the end, as the high contributions they have observed previously may lead them to expect more of the same.

will not be affected by concern withdrawal, since those who misbehave are not in their group for very long. Furthermore, a non-myopic free-rider may well cooperate strategically, as it rapidly becomes apparent that contributing nothing will very likely lead to being unmatched (or matched involuntarily with other undesirables), with an average payoff near 25.[30] Even if one has selfish preferences and understands the theoretical equilibrium if everyone has such preferences, own-payoff-maximization may require full contribution, given expectations that a sufficient number of others will contribute (and this is observed after the first round, in any case).

People base their current actions on what they expect the future to be, and particularly on how their own actions may affect this future. The precise conditions under which a purely selfish individual should choose cooperation over free-riding will depend on beliefs, but cooperation will not be so rare given the payoff attraction of large groups.[31] The outlook of the efficiency potential of large groups, combined with the expectation of such large groups forming, makes current risky actions worthwhile and cooperation is sufficiently safe-guarded by the endogenously severe impact of exclusion. Thus, we see behavior in our setting as being primarily determined by a combination of altruism and non-myopic own-payoff maximization.

In this regard, an important issue is the (perceived) opportunity cost of misbehavior. The outlook of expulsion from the grand coalition in Treatment 2 is certainly grim, comparable to that in Cinyabuguma *et al.* (2005). This is, however, a desired impact from our endogenous

---

[30] Of course, this depends on the presence of a 'critical mass' of cooperators in the population, as a population comprised of only free-riders will never get off the ground. But our data for the three-person groups in the first round shows consistently high (74-81%) contribution rates in all three treatments.

[31] For example, consider a person with selfish preferences in an n-person group at the beginning of VP4, who is pondering whether to free ride or to stick to the norm. Free riding gets him expelled so that he is likely to become a single or in a forced group of free-riders. Thus his average per period payoff in this case is: {25*f(n)*[n-1]/n + 25} +25. If he postpones free-riding to VP5 while sticking to the norm of full contribution in VP4, then he expects payoff {25*n*f(n)/n} + { z*25*[n-1]*f(n)/n +25}, where z denotes the aggregate % of contribution from other members in the group. For the subject to decide against free-riding in VP4, we need {25*f(n)*[n-1]/n + 25} +25 < {25*n*f(n)/n} + { z*25*[n-1]*f(n)/n +25}. In Treatment 2, z must be greater than 31% (50%) if he is in a 9-person (4-person) group. In Treatment 3, z must be greater than 57% (50%) if he is in a 9-person (4-person) group. There is correspondingly greater incentive to stick to the full-contribution norm within large groups in earlier rounds.

group-formation setup, where the cost of exclusion varies. In a society with smaller groups, the opportunity cost is low, so that people may not care much about being excluded.[32] However, with a dominant large group in place, the consequence of being excluded is more dramatic, and thus so is the disciplinary impact induced by this measure. So, in some sense, our design endogenously induces a condition that maximizes the disciplinary effect of exclusion.[33]

We must also explain the unraveling in the last round of each segment of Treatments 2 and 3. One clue is that if we look at each of the three periods in the last round, we find that in fact, the unraveling increases dramatically over the course of this round. In the first segment, in period 13 the proportion of zero contributions was .03; this increased to .11 in period 14, and to .21 in period 15. In the second segment, in period 28 the proportion of zero contributions was .20; this increased to .27 in period 29, and to .46 in period 30.

Our explanation is that some of the free-riders who have been masquerading as cooperators realize in the first period of the last round that there is no longer *any* strategic reason to make positive contributions, and so they stop doing so. This leads to the light progressively dawning on other free-riders, and conditional cooperators abandoning cooperation in the face of increasingly selfish behavior. The higher zero-contribution rate in the second segment reflects learning that it is not uncommon for people to stop contributing in the last round of the segment.

5. Conclusion

We provide an assortative-matching mechanism in a public-goods environment, where the issue of how to implement collective action in the face of potential free-riding is a crucial

[32] Also, the endogenous impact of exclusion is less severe initially, as smaller groups have low production efficiency in any case, while one might still have a possibility of joining a large, more efficient and rewarding group.
[33] A similar dynamic effect is also implicit in Gürerk, Irlenbusch, and Rockenbach (2006) in the context of competition between institutions with feasible sanctions and sanction-free institutions.

one. We find a great degree of success for this mechanism, with the average contribution rate quite high in two treatments with endogenous exit, exclusion, and merger; this contribution rate is slightly (but significantly) smaller in the second treatment, where we cap the social value of a contribution, so that there is no efficiency advantage to being in a large group. In the first treatment, we see large and stable groups forming, but there is considerably more instability and smaller group sizes in the second treatment.

People are often conditional cooperators. They need environments or institutional setups that feature favorable dynamics to 'rein them in' and to help them feel comfortable with cooperative behavior. Institutions facilitate proliferation of cooperation but do not necessarily guarantee it.[34] Our design offers safeguards that make it difficult for myopic free-riders to seriously interfere with the endeavor of smart "conditional cooperators". The driving force appears to be the economies of scale combined with the awareness that bad behavior will result in ostracism, in the Athenian sense including possible redemption. This redemption is a unique feature of our environment, applying to about one-third of our population.

We are not the first to achieve high cooperation rates through exclusion; consider the results in Cinyabuguma *et al.* (2005). However, note that in addition to featuring permanent exclusion, they *began* with the grand coalition, rather than having to create it from small groups. Thus, there is no possibility of increasing group size with their design. Furthermore, their form of ostracism is quite draconian. This feature does not seem a particularly apt depiction of the field, as permanent exclusion seems unrealistic and unattractive. Exclusion is important – witness the trend towards gated communities and security buildings – but the notion of redemption is also attractive, and here applies to one-third of the population.

[34] See Yang, Yue, and Yu (2007) for such a case in an assortative matching prisoners' dilemma experiment.

We also feel it is important for established groups to have a possibility of joining together for their mutual benefit.[35] The speed at which groups can grow is a relevant issue. In our design, mergers serve as a useful complement to the scale economy and we find that cooperative groups can grow quickly. Our results are in contrast with those in Weber (2005, 2006), where managers can induce a high rate of contribution in the minimum-effort game with large groups by slowly building these groups, but managers who try to increase group size rapidly have a notable lack of success in achieving good outcomes in the larger groups. We suspect that the lack of scale economies in the minimum-effort game make it much more difficult to sustain coordination in larger groups, as there is no real reason to form larger groups if one can achieve the highest possible payoff with the relatively easy coordination in small groups. Our flexible design plus economies of scale might yield interesting extensions to this approach as well.

Ours is an approximation of how groups (and groups of groups) can evolve, adapted so that it is manageable for conducting experiments while preserving the essential features. The mechanism differs from those in previous studies with respect to being fluid and dynamic; it is the first to provide a general and flexible platform to address both increases and decreases in endogenous size. Change is certainly a prominent feature of our contemporary world, and allowing both inflows and outflows between and among groups is a better representation of the environment in the field.

While we have established that an assortative mechanism *can* work, we have not yet established the threshold conditions that must be satisfied; in Treatment 3 we have begun the investigation of the boundaries for our institution. Different specifications regarding the rules of voting, the production function for public goods, the information available, and the degree of

[35] One can consider the background context to be legally-binding agreements in business and life (as in Hauk and Nagel 2001).

friction present may cause completely different inner dynamics of evolution of cooperation. One critical element is clearly the greater efficiency in larger groups, while a second critical element is the insurance that would-be cooperators have that they will be able to ward off would-be free-riders. In future work, we will explore the boundary of cooperation within this general setting.

References

Ahn, T., R. Isaac, and T. Salmon (2006), "Endogenous Group Formation," mimeo.

Andreoni, J. and L. Samuelson (2006), "Building Rational Cooperation," *Journal of Economic Theory*, **127**, 117-154.

Bergstrom, T. (2002), "Evolution of Social Behavior: Individual and Group Selection," *Journal of Economic Perspectives*, **16**, 67-88.

Brosig, J., M. Margreiter, and J. Weimann (2005), "Endogenous Group Formation and the Provision of Public Goods: The Role of Promises and Lies," mimeo.

Charness, G. and M. Rabin (2002), "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, **117**, 817-969.

Cinyabuguma, M., T. Page, and L. Putterman (2005), "Cooperation under the Threat of Expulsion in a Public Goods Experiment," *Journal of Public Economics*, **89**, 1421-1435.

Coricelli, G., D. Fehr, and G. Fellner (2005), "Partner Selection in Public Goods Experiments," mimeo.

Croson, R. (2001), "Feedback in voluntary contribution mechanisms: An experiment in team production," *Research in Experimental Economics*, **8**, 85-97.

Croson, R., E. Fatás, and T. Neugebauer (2006), "Excludability and Contribution: A Laboratory Study in Team Production," mimeo.

Cuellar, A., L. McReynolds, and G. Wasserman (2006), "A cure for crime: Can mental health treatment diversion reduce crime among youth?," *Journal of Policy Analysis and Management*, **25**, 197-214.

de Waal, F (1986), "The brutal elimination of a rival among captive male chimpanzees," *Ethology and Sociobiology*, **7**, 237-251.

Ehrhart, K.-M. and C. Keser (1999), "Mobility and Cooperation: On the Run," mimeo.

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

Fehr, E. A. Klein, and K. Schmidt (2007), "Fairness and Contract Design," *Econometrica*, **75**, 121-154.

Fehr, E. and B. Rockenbach (2003), "Detrimental Effects of Sanctions on Human Altruism," *Nature*, **422**, 137-140.

Fischbacher, U. and S. Gächter (2006), "Heterogeneous Social Preferences and the Dynamics of Free Riding in Public Goods," mimeo.

Frank, R. (1988), *Passions within Reason: The Strategic Role of the Emotions*, New York: W.W. Norton.

Goodall, J. (1986), "Social Rejection, Exclusion, and Shunning among the Gombe Chimpanzees," *Ethology and Sociobiology*, **7**, 227-236.

Gruter, M. and R. Masters (1986), "Ostracism as a social and biological phenomenon: An introduction," *Ethology and Sociobiology*, **7**, 149-158.
Gunnthorsdöttir, A., D. Houser, and K. McCabe (2007): "Disposition, History and Contributions in Public Goods Experiments", *Journal of Economic Behavior and Organization*, **62**, 304-315.
Gürerk, Ö., B. Irlenbusch, and B. Rockenbach (2006), "The Competitive Advantage of Sanctioning Institutions," *Science*, **312**, 108-111.
Hauk, E. and R. Nagel (2001), "Choice of Partners in Multiple Two-Person Prisoner's Dilemma Games," *Journal of Conflict Resolution*, **6**, 770-793.
Lancaster, J. (1986), "Primate Social Behavior and Ostracism," *Ethology and Sociobiology*, **7**, 215-225.
Masclet, D., C. Noussair, S. Tucker, and M.-C. Villeval (2003), "Monetary and Non-Monetary Punishment in the Voluntary Contribution Mechanism, *American Economic Review*, **93**, 366-380.
Maier-Rigaud, F., P. Martinsson, and G. Staffiero (2005), "Ostracism and the Provision of a Public Good: Experimental Evidence," mimeo.
Marwell, G. and R. Ames (1979), "Experiments on the Provision of Public Goods. **I**. Resources, Interest, Group Size, and the Free Rider Problems," *American Journal of Sociology*, **84**, 1335-1360.
Ohtsuki, H., and Y. Iwasa (2006), "The leading eight: Social norms that can maintain cooperation by indirect reciprocity," *Journal of Theoretical Biology*, 239, 435-444.
Önes, U. and L. Putterman (2007), "The Ecology of Collective Action: A Public Goods and Sanctions Experiment with Controlled Group Formation," *Journal of Economic Behavior and Organization*, **62**, 495-521.
Page, T., L. Putterman, and B. Unel (2005), "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry, and Efficiency", *Economic Journal*, 115, 1032-1053.
Riedl, A. and A. Ule (2002), "Exclusion and Cooperation in Social Network Experiments," mimeo.
Sethi, R. and E. Somanathan, "Understanding Reciprocity," *Journal of Economic Behavior and Organization*, **50**, 1-27.
Siegel, S. and N. Castellan (1988), *Nonparametric Statistics for the Behavioral Sciences*, Boston: McGraw-Hill.
Sober, E. and D. Wilson (1998), *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge, MA: Harvard University Press.
Tiebout, C. (1956), "A Pure Theory of Local Expenditures," *Journal of Political Economy*, **64**, 416-424.
Weber, R. (2005), "Managing Growth to Achieve Efficient Coordination in Large Groups: Theory and Experimental Evidence," mimeo.
Weber, R. (2006), "Managing Growth to Achieve Efficient Coordination in Large Groups," *American Economic Review*, **96**, 114-126.
Williams, K. (2001), *Ostracism: The Power of Silence*, New York: Guilford Publications.
Yang, C.-L., C.-S. Yue and I-T. Yu (2007), "The Rise of Cooperation in Correlated Matching Prisoners Dilemma: An Experiment," *Experimental Economics*, 10, 3-20.

Appendix A

INSTRUCTIONS

Welcome to our experiment. You will receive $5 for showing up, in addition to your earnings from the session. There are 18 participants in the session. These participants will be subdivided into two *societies* of nine people each, with one's society determined by a random draw. You will not know which participants are in your society.

There will be a total of 30 periods in the session, comprised of two segments of 15 periods each. In each period, you will be in a *group*. All groups once formed remain intact for three periods. For the first 3 periods of each 15-period segment, we will randomly form 3 groups of 3 participants within each 9-person society (these societies will not be the same in the 2 different segments.) After this, groups will be formed based on the votes of the participants in the society.

Each individual's has a unique identification number over a 15-period segment; this identification number is randomly re-drawn in the second 15-period segment.

Once the groups are formed, each person will receive 25 tokens in each period and will make choices regarding individual and group allocations. The choice is how many tokens to allocate to the *group account* and how many to allocate to one's *private account*. Each token allocated to one's private account pays one unit to that individual, independently of the group size. On the other hand, each token allocated to the group account yields a return higher than one unit and all members of the group share equally the returns from the tokens allocated to the group account. Moreover, the rate (multiplying factor) of return increases with the group size. The table below shows the factor for each group size:

n	1	2	3	4	5	6	7	8	9
Factor	1	1.322	1.455	1.600	1.760	1.936	2.130	2.343	2.577
Per-person return	1	0.661	0.485	0.400	0.352	0.323	0.304	0.293	0.286

Thus, for the same amount of tokens allocated to the group account, larger groups will earn more money. On the other hand, the per-person return from the same amount allocated to the group account diminishes as the group size increases.

So, for example, if in a **four-person group** the total number of tokens in the group account is 10, then the total return to the group from the group account is 16, and each person receives a payoff of 4 units from the group account. For every additional token invested in the group account, everyone receives an additional return of 0.4 units (1.600/4), while whoever makes the additional contribution will have one unit less available for his/her private account.

By comparison, if in a **seven-person group** the total number of tokens in the group account is 10, then the total return to the group from the group account is 21.3, and each person receives a payoff of 3.04 units from the group account. For every additional token invested in the group account, everyone receives an additional return of 0.304 units (2.130/7), while whoever makes the additional contribution will have one unit less available for his/her private account.

In the beginning of the 15-period segment, each person is initially assigned to a three-person group. There are three periods of decisions made about the allocation of tokens before people vote on possible changes of groups. After each period, each person learns only his or her group's total (average) contribution level to the group account.

As mentioned above, groups remain intact for three periods and are then re-formed for the next three periods. We first allow each person to exit voluntarily from his or her group, and then the people remaining in each group decide if they wish to exclude any remaining group members from the group. Then, mergers among the group remnants and voluntary and involuntary singles will be considered. Here is how the process of group formation works:

Stage 1. Prior to voting, each person is told the average contribution level for every individual. in his or her *own* group, and the average for every existing group, during the past 3 periods. Each person then chooses whether he or she wishes to stay in his or her current group. If the answer is "no", he or she will be considered a "single" temporarily and will not participate in the following voting among those who chose "yes". If the answer is "yes", then this person casts a vote with respect to each other member of the group, regarding whether to keep the other person in the group. After this voting round, those persons who choose to stay in the group and who receive at least 50% *positive* votes from the other willing-to-stay group members remain in the group. All others are considered "singles" temporarily.

Stage 2. At this point, we have a number of groups (group remnants) and perhaps a number of single participants not in groups. Each person will now be told the average contribution for the past three periods for each of these temporary entities (groups or singles), as well as the composition (identification numbers) of these entities. For mergers among those entities, now each person casts a vote about **every other** entity in his or her society, expressing whether he or she would like his or her current group (or single) to merge with each other entity (group or single).

Stage 3. Groups are now re-formed based on these votes. We rank the groups remaining after Stage 1 by size, with random ordering for groups with the same size. Next, we check to see whether the largest group wishes to merge with the next-largest group, which may also be a "single". If we do not have *mutual approval* of this merger, we check whether the largest group and the third-largest group wish to merge etc. till one such merger takes place. If there are no such mergers with this largest group, we check the next-largest group in combination with the smaller groups. *Whenever there is a merger, we repeat the same procedure starting with the now-largest group*. We continue this process until we cannot find two groups (singles included) that wish to merge.

For *mutual approval* of a merger, we require that, <u>for each group</u>, at least 60% of the votes are in favor of merging with the other group. Note however that all votes are carried over after each merger as illustrated in the following example. Suppose there are three intermediate groups A, B, and C of sizes 3, 2, and 1. The group A's votes for mergers with B and C are 1/3 and 2/3 respectively. B's for A is 2/2, for C is 1/2. C's are 1/1 for A and 1/1 for B. According to our procedure, the first merger is A with C. This new intermediate entity AC's vote for B, now, is 2/4, while B's for AC becomes 2/2*3/4 + 1/2*1/4 = 7/8 as this new entity AC is composed to 3/4 of the old A and to 1/4 of the old C.

There may be single participants remaining at the end of this process. If there is more than one such single participant remaining, these participants will be randomly grouped into 3-person and 2-person groups, in group formats of 3-3-3, 3-3-2, 3-2-2, 3-3, 3-2, 2-2, 3, 2, if there are 9, 8, 7, 6, 5, 4, 3, or 2 singles, respectively. If only one is left, he will be a single through the next 3 rounds.

At the end of the experiment, we will add up your payoffs from all 30 periods and convert them to actual dollars at the rate of $0.01 for each unit. We will then pay each participant individually and privately. Thank you again for your participation in our research.

Appendix B

Table B1: Individual contribution and redemption in Treatment 2

ID	VP 1	VP 2	VP 3	VP 4	VP 5	VP 6	VP 7	VP 8	VP 9	VP 10	Redeemed
1	7.3	**19.0**	23.7	19.0	***25.0***	25.0	25.0	23.0	***0.0***	***0.0***	1
2	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	16.7	
3	13.3	21.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	1
4	10.0	21.7	**22.3**	25.0	25.0	25.0	25.0	25.0	25.0	16.7	1
5	12.3	12.7	**12.7**	*25.0*	25.0	24.0	24.7	24.3	25.0	25.0	1
6	11.7	**18.0**	24.0	25.0	25.0	25.0	25.0	24.3	25.0	25.0	1
7	8.3	25.0	25.0	25.0	25.0	24.3	25.0	24.7	25.0	23.7	1
8	16.7	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	
9	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	
10	13.0	**7.7**	7.7	**10.3**	**5.7**	16.0	***24.7***	***25.0***	*8.3*	**8.3**	
11	20.0	**23.0**	***25.0***	25.0	**25.0**	25.0	25.0	25.0	25.0	13.3	
12	25.0	25.0	25.0	25.0	16.7	25.0	25.0	25.0	25.0	0.0	
13	13.0	**21.7**	25.0	25.0	25.0	25.0	25.0	25.0	25.0	2.7	1
14	21.7	25.0	25.0	25.0	16.7	25.0	25.0	25.0	25.0	16.7	
15	20.3	25.0	25.0	25.0	25.0	25.0	25.0	25.0	6.7	**25.0**	
16	25.0	25.0	25.0	25.0	20.0	25.0	25.0	25.0	25.0	25.0	
17	13.7	**15.3**	**17.0**	***25.0***	24.7	23.3	25.0	25.0	25.0	15.0	1
18	21.7	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	8.3	
19	16.0	**20.7**	**25.0**	25.0	25.0	25.0	25.0	25.0	25.0	25.0	
20	18.3	**25.0**	25.0	25.0	25.0	25.0	25.0	25.0	25.0	8.3	
21	4.7	**18.3**	20.7	**23.3**	**25.0**	25.0	25.0	25.0	25.0	25.0	1
22	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	
23	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	
24	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	8.3	
25	20.3	25.0	25.0	25.0	8.3	25.0	25.0	25.0	25.0	8.3	
26	25.0	25.0	25.0	25.0	0.0	25.0	25.0	25.0	25.0	0.0	
27	21.7	**25.0**	25.0	25.0	25.0	25.0	25.0	25.0	25.0	21.7	
28	23.3	25.0	25.0	25.0	16.7	25.0	25.0	25.0	25.0	8.3	
29	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	0.0	0.0	
30	10.0	16.7	19.0	25.0	25.0	6.7	**8.3**	**8.3**	**1.7**	**3.3**	
31	15.0	**19.3**	20.0	22.3	25.0	24.7	25.0	25.0	25.0	25.0	1
32	13.3	**16.7**	**20.0**	**21.3**	***23.3***	11.7	**12.0**	**3.3**	**2.7**	**2.7**	
33	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	0.0	
34	25.0	25.0	25.0	25.0	16.7	25.0	25.0	25.0	25.0	1.7	
35	9.3	16.0	21.7	25.0	25.0	24.0	25.0	25.0	25.0	16.3	1
36	23.3	24.3	25.0	25.0	11.7	25.0	25.0	25.0	25.0	11.7	
37	19.3	25.0	25.0	25.0	25.0	23.3	25.0	25.0	16.7	**0.0**	
38	13.0	**13.0**	**6.3**	**4.0**	3.0	4.3	**5.3**	**2.7**	**2.7**	**3.0**	
39	18.7	**18.3**	20.0	24.0	25.0	21.7	25.0	25.0	25.0	25.0	
40	21.3	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	8.3	
41	23.3	25.0	25.0	25.0	25.0	24.0	25.0	25.0	25.0	3.3	
42	13.0	**15.0**	**17.7**	**24.7**	20.7	22.7	24.7	**24.3**	25.0	22.7	1
43	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	16.7	

44	18.3	24.0	25.0	25.0	19.3	20.7	**25.0**	**25.0**	**25.0**	25.0	
45	7.7	**11.3**	**6.0**	**11.7**	10.0	17.7	**21.7**	**24.7**	**25.0**	24.0	1
46	23.3	25.0	25.0	25.0	3.3	25.0	25.0	25.0	25.0	12.7	
47	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	
48	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	
49	14.0	**20.3**	18.0	***25.0***	25.0	25.0	25.0	25.0	25.0	16.7	1
50	21.7	25.0	25.0	25.0	8.3	25.0	25.0	25.0	25.0	10.0	
51	15.0	**18.3**	**23.3**	25.0	16.7	25.0	25.0	25.0	25.0	8.3	1
52	16.7	**21.7**	25.0	25.0	25.0	25.0	25.0	25.0	25.0	0.0	
53	25.0	25.0	25.0	25.0	0.0	25.0	25.0	25.0	25.0	0.0	
54	15.0	19.0	**25.0**	25.0	25.0	25.0	25.0	25.0	25.0	16.7	1
55	25.0	25.0	25.0	25.0	16.7	25.0	25.0	25.0	25.0	0.0	
56	18.0	22.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	4.0	
57	20.0	**25.0**	25.0	25.0	25.0	25.0	25.0	25.0	25.0	21.7	
58	21.7	**21.7**	**21.7**	***25.0***	23.3	23.3	25.0	25.0	25.0	23.3	
59	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	24.7	
60	15.0	21.3	25.0	25.0	8.0	24.7	25.0	25.0	25.0	0.0	1
61	15.0	19.7	**20.7**	23.0	20.7	18.3	***22.3***	*25.0*	22.3	***25.0***	1
62	25.0	25.0	25.0	25.0	16.7	25.0	25.0	25.0	25.0	16.7	
63	15.0	**23.3**	25.0	25.0	9.0	25.0	25.0	25.0	25.0	0.0	1

Note: Italics mean the person was a single during the VP; bold means it was an exogenous group.

Table B2: Individual contribution and redemption in Treatment 3

ID	VP 1	VP 2	VP 3	VP 4	VP 5	VP 6	VP 7	VP 8	VP 9	VP 10	Redeemed
64	21.7	21.7	21.7	20.0	14.3	21.7	**21.0**	**21.7**	**21.0**	**18.3**	
65	19.7	20.7	22.0	23.7	24.0	21.7	**25.0**	25.0	24.7	25.0	
66	6.0	**5.0**	**5.0**	**16.7**	21.7	21.7	18.3	**23.3**	**23.3**	7.0	1
67	11.3	23.3	25.0	23.3	25.0	24.0	25.0	25.0	25.0	16.7	1
68	6.0	**7.0**	**10.7**	**11.0**	15.0	22.3	**20.0**	**25.0**	25.0	16.7	1
69	17.7	19.0	19.7	19.7	***6.7***	19.0	**23.3**	19.7	**20.0**	0.0	
70	16.7	**12.7**	**2.0**	**10.0**	23.3	21.0	**25.0**	25.0	8.3	**0.0**	1
71	18.3	23.3	25.0	24.0	25.0	24.3	25.0	25.0	25.0	25.0	
72	22.7	25.0	25.0	25.0	25.0	23.3	16.7	**25.0**	25.0	25.0	
73	4.7	***15.3***	***21.7***	***23.3***	16.7	16.0	**21.3**	**23.0**	22.3	23.0	1
74	20.0	20.0	22.0	24.0	20.0	23.0	24.3	25.0	25.0	19.7	
75	20.0	23.3	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	
76	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	0.0	
77	21.7	18.3	21.7	21.7	20.0	18.3	**20.0**	**21.7**	**20.0**	6.7	
78	25.0	25.0	24.0	23.7	***0.0***	25.0	25.0	25.0	25.0	16.7	
79	10.3	19.3	25.0	25.0	25.0	24.3	24.3	25.0	21.7	**21.7**	1
80	24.7	21.7	22.0	24.0	24.3	22.7	24.7	24.3	24.7	24.3	
81	23.3	25.0	24.7	25.0	25.0	24.7	24.3	25.0	24.7	25.0	
82	11.7	**25.0**	18.3	**25.0**	24.3	20.0	20.0	21.7	25.0	16.7	1
83	19.0	20.3	18.3	**22.3**	24.0	24.0	21.3	22.7	23.3	23.0	
84	20.0	25.0	24.3	24.7	16.7	22.0	25.0	25.0	25.0	16.7	
85	21.0	20.0	24.3	23.7	23.3	21.7	25.0	22.7	23.3	15.0	
86	25.0	25.0	25.0	25.0	25.0	25.0	21.7	**16.7**	**25.0**	**18.3**	

87	13.3	**18.3**	***24.0***	18.3	**20.0**	18.3	**15.0**	**7.0**	**18.3**	**10.0**	
88	20.0	21.3	22.7	23.7	17.3	22.0	23.0	23.3	23.3	11.7	
89	15.0	23.0	23.7	25.0	0.0	22.3	25.0	24.0	24.7	**0.0**	1
90	22.7	25.0	25.0	24.7	**24.0**	25.0	25.0	25.0	22.0	**6.7**	
91	14.0	15.3	**17.0**	15.0	**16.7**	13.0	***22.3***	17.3	**15.7**	**10.0**	
92	5.7	**19.0**	21.7	18.3	**19.3**	19.0	16.7	21.3	**15.0**	**11.7**	1
93	16.7	14.0	**17.7**	11.3	**13.3**	13.7	**8.7**	**1.0**	**20.0**	**15.7**	
94	21.7	23.3	23.3	23.3	0.0	20.7	25.0	***25.0***	24.3	0.0	
95	9.3	17.7	20.3	21.7	8.3	20.0	25.0	25.0	25.0	25.0	1
96	11.7	17.0	18.7	20.0	21.0	20.0	19.7	**24.0**	17.7	**13.0**	1
97	25.0	23.3	20.0	21.7	15.0	24.3	22.0	22.3	21.7	0.0	
98	10.0	16.7	20.0	22.0	5.0	21.0	23.0	22.0	0.0	**0.0**	1
99	8.3	**20.0**	23.3	23.3	13.3	21.7	25.0	21.7	23.3	**23.3**	1
100	12.3	**9.7**	**14.3**	**16.3**	10.0	19.7	**24.3**	24.0	**15.0**	12.3	1
101	10.7	**13.3**	**20.0**	**14.0**	**19.7**	19.7	**20.0**	***25.0***	20.0	21.0	1
102	25.0	25.0	25.0	24.7	25.0	24.3	**16.7**	**16.7**	**25.0**	16.7	
103	21.7	23.3	24.3	25.0	25.0	25.0	25.0	25.0	25.0	20.0	
104	16.7	20.7	23.3	24.3	25.0	16.7	**18.3**	21.7	15.0	15.0	
105	11.7	**21.7**	25.0	21.7	25.0	25.0	25.0	25.0	25.0	20.0	1
106	16.7	**16.0**	15.3	**15.0**	**18.0**	23.7	25.0	25.0	25.0	16.7	1
107	8.7	**17.3**	**14.7**	**16.7**	**19.3**	25.0	25.0	25.0	25.0	25.0	1
108	20.3	20.7	22.3	**18.7**	**18.0**	25.0	25.0	24.7	17.3	**22.3**	
109	14.0	12.7	**23.3**	**20.0**	**13.3**	25.0	25.0	25.0	25.0	16.7	1
110	10.0	14.7	**17.3**	**19.0**	**17.7**	15.7	**20.0**	21.0	21.7	**14.3**	1
111	21.7	25.0	25.0	25.0	16.7	23.3	**23.3**	25.0	25.0	11.7	
112	25.0	25.0	25.0	25.0	8.3	25.0	25.0	25.0	25.0	8.3	
113	21.7	**25.0**	25.0	25.0	8.3	25.0	25.0	25.0	25.0	16.7	
114	23.3	25.0	25.0	25.0	16.7	25.0	25.0	25.0	21.7	11.7	
115	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.0	**16.7**	**25.0**	
116	15.0	**21.7**	21.7	16.7	**15.3**	17.3	**15.0**	18.7	**16.7**	**22.0**	
117	16.0	**21.7**	25.0	25.0	16.7	13.3	**15.0**	**13.3**	**10.0**	**5.0**	

Note: Italics mean the person was a single during the VP; bold means it was an exogenous group.

Appendix C

Table C1: Group sizes and mergers over time, by Session, Society, Segment

	VP1	VP2	VP3	VP4	VP5
Treatment 2					
1,1,1	3-3-3	*5-2-2*	5-2-2	*8-1*	8-1
1,1,2	3-3-3	**9-0**	9-0	8-1	8-1
2,1,1	3-3-3	4-3-2	*6-2-1*	7-2	7-2
2,2,1	3-3-3	*6-3*	7-2	8-1	9-0
2,1,2	3-3-3	*6-3*	*9-0*	9-0	9-0
2,2,2	3-3-3	**8-1**	8-1	8-1	7-2
3,1,1	3-3-3	*5-2-2*	5-2-2	5-2-2	6-2-1
3,2,1	3-3-3	*5-2-2*	5-2-2	*7-2*	7-2
3,1,2	3-3-3	**7-2**	6-3	7-2	*7-2*
3,2,2	3-3-3	6-3	6-3	6-3	6-3
4,1,1	3-3-3	3-2-2-2	**7-2**	8-1	9-0
4,2,1	3-3-3	*5-2-2*	*7-2*	8-1	9-0
4,1,2	3-3-3	**9-0**	9-0	9-0	9-0
4,2,2	3-3-3	**8-1**	8-1	9-0	8-1
Treatment 3					
5,1,1	3-3-3	4-3-2	4-3-2	*6-3*	4-2-2-1
5,2,1	3-3-3	*6-2-1*	5-3-1	*8-1*	6-2-1
5,1,2	3-3-3	5-2-2	*5-2-2*	5-2-2	3-2-2-2
5,2,2	3-3-3	4-3-2	4-3-2	4-3-2	4-3-2
6,1,1	3-3-3	*4-3-2*	4-4-1	4-3-2	3-2-2-2
6,2,1	3-3-3	*5-2-2*	*5-2-2*	5-2-2	4-3-2
6,1,2	3-3-3	**8-1**	6-2-1	7-2	5-2-2
6,2,2	3-3-3	4-3-2	3-2-2-2	3-2-2-2	3-3-3
7,1,1	3-3-3	*4-3-2*	4-3-2	4-3-2	3-2-2-2
7,2,1	3-3-3	4-3-2	4-3-2	4-3-2	4-3-2
7,1,2	3-3-3	3-3-3	*5-2-2*	4-3-2	5-2-2
7,2,2	3-3-3	4-3-2	4-4-1	4-3-2	4-3-2

Entries show the groups existing at the start of the voting period. Entries in italics mean that there was one merge during the previous voting period; entries in bold mean that there were two merges.

Appendix D

Figure D1: Average contributions and profits in Treatment 2, by individual

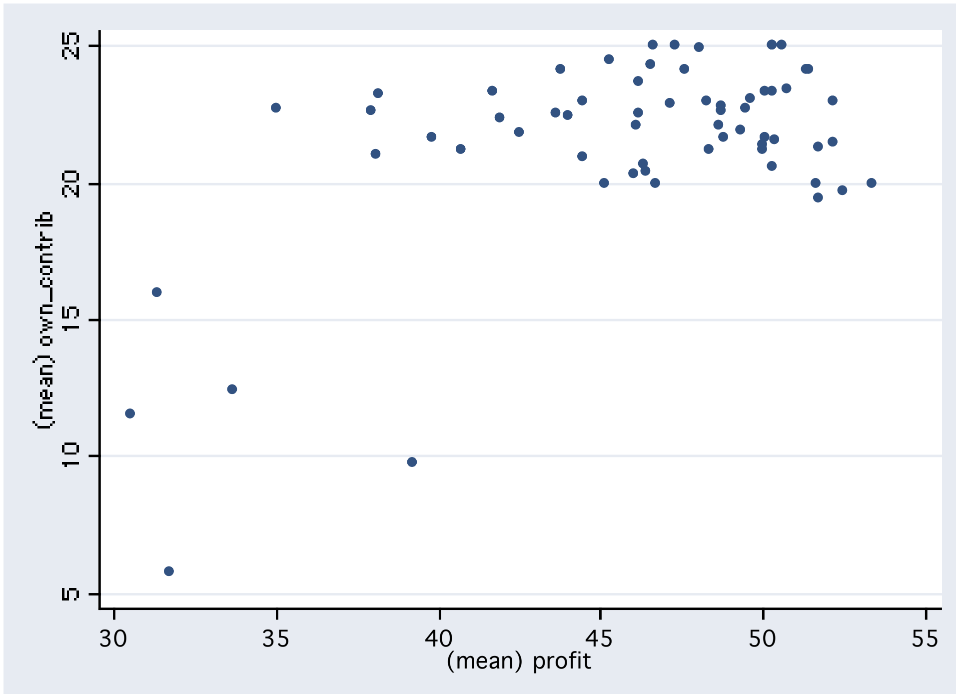


Figure D2: Average contributions and profits in Treatment 3, by individual