



Sustaining cooperation through self-sorting: The good, the bad, and the conditional

Karen Evelyn Hauge^{a,1}, Kjell Arne Brekke^{b,1}, Karine Nyborg^{b,1,2}, and Jo Thori Lind^{b,1}

^aRagnar Frisch Centre for Economic Research, N-0347 Oslo, Norway; and ^bDepartment of Economics, University of Oslo, N-0033 Oslo, Norway

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved June 28, 2018 (received for review March 1, 2018)

In four public-good game experiments, we study self-sorting as a means to facilitate cooperation in groups. When individuals can choose to join groups precommitted to charity, such groups sustain cooperation toward the group's local public good. By eliciting subjects' conditional contribution profiles, we find that subjects who prefer the charity groups have higher average conditional contribution levels but do not differ with respect to the slope of their profiles. The majority of subjects in both group types are conditional cooperators whose willingness to contribute is stimulated by generous group members but undermined by free-riders. Charity groups thus seem better able to sustain cooperation because they attract a greater number of more generous individuals, triggering generous responses by conditional cooperators.

voluntary contributions | altruism | conditional cooperation | corporate social responsibility | public goods

Environmental problems are often associated with social dilemmas: While individual contributions would benefit the group at large, contributing is individually unprofitable (1). Similar issues arise with teamwork: If individual efforts are not monitored, efforts are essentially voluntary contributions to the team's production (2). The social bonds and well-being among a group of friends, acquaintances, or neighbors depend on parties' willingness to invest time and effort in their community, e.g., by organizing social events, helping others, or maintaining community property (3).

In some cases, formal regulation can be used, e.g., performance pay, direct orders, subsidies and green taxes, tradable and nontradable emission permits. Nevertheless, for various reasons, e.g., limited observability, formal regulation is not always feasible or effective. In such cases, voluntary contributions may be essential.

Even when purely self-interested agents have few incentives to cooperate, there is substantial empirical evidence that groups sometimes do manage to sustain cooperation (4). Nevertheless, as demonstrated in numerous economic experiments using the public good game, cooperation tends to crumble over time unless formal or informal sanctioning mechanisms are present or individuals are matched in groups so that the more cooperative individuals interact mostly with one other (5, 6).

An important explanation for the latter result is that many experimental subjects are conditional cooperators who are willing to contribute more when others contribute more (7–14). When interacting with others who contribute high amounts, these subjects are willing to contribute a lot; if matched with free-riders, however, they reduce or stop their contributions. Consequently, if subjects are matched exogenously according to their previous contribution levels, or if individuals can influence group composition based on observed contributions, groups with initially high contributions often succeed in sustaining their cooperation (15–21).

Such sorting mechanisms cannot always be used, however; for example, information about previous contribution levels may not be available. It is thus of interest to explore other possible mechanisms for matching high contributors. The present paper studies a mechanism based on self-selection that does not rely on observed previous behaviors.

In ref. 22, referred to below as “study 1,” we studied a public good game in which some groups (“red groups”) were precommitted to

donate a fixed amount to a charity, while other groups (“blue groups”) were not. We found that contributions to the group's local public good were higher in red groups than in blue groups; moreover, while blue groups displayed the usual pattern of declining contributions over time (5), red groups sustained their cooperation.

In the present paper, we present additional experimental evidence (studies 2–4) exploring the mechanisms behind the higher and sustained contributions in red groups. These three studies replicate study 1, although with slight variations. Most importantly, these studies provide information about subjects' conditional contribution profiles, that is, their willingness to contribute given the contribution levels of other group members. We explore how conditional contribution profiles differ in individuals who self-select into red and blue groups and whether these differences can explain the dynamics allowing sustained cooperation in red but not in blue groups.

Below, we use two approaches to describe conditional contribution profiles. First, we classify all subjects as altruists, free-riders, conditional cooperators, or others (8, 23); second, we use information on the slope and average level of each subject's conditional contribution profile (24).

In particular, we explore (*i*) whether subjects classified as altruists, free-riders, conditional cooperators, and others differ in their propensity to self-select into red groups and (*ii*) whether the slope and/or average contribution level of subjects' conditional contribution profiles are associated with self-selection into red groups.

Our main finding is that more generous subjects—altruists and, more generally, subjects with high average contribution levels—self-select into red groups, while the less generous tend to choose blue groups. Since the large majority of our subjects are conditional cooperators, this provides an explanation of the findings of ref. 22 that cooperation is sustained over time in red groups but not in blue groups: Subjects with more generous conditional contribution profiles support the good will of conditional cooperators in red groups, while less generous subjects undermine it in blue groups.

However, we do not observe significant differences in the slope of conditional contribution profiles between the red and blue groups.

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, “Economics, Environment, and Sustainable Development,” held January 17–18, 2018, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/economics-environment-and.

Author contributions: K.E.H., K.A.B., K.N., and J.T.L. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the [PNAS license](http://www.pnas.org/licenses).

Data deposition: The data reported in this paper have been deposited in the Harvard Dataverse (<https://doi.org/10.7910/DVNU/US270>).

¹K.E.H., K.A.B., K.N., and J.T.L. contributed equally to this article.

²To whom correspondence should be addressed. Email: karine.nyborg@econ.uio.no.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1802875115/-DCSupplemental.

Published online July 30, 2018.

Sustained cooperation in red groups but not in blue groups might also be due to changed preferences or self-selection patterns over time or to strategic considerations. We find little support for these alternative hypotheses.

Experimental Design

We report on four studies, all of which are based on variations of a standard linear public good game. Study 1 is based on data from ref. 22, but studies 2–4 are previously unpublished.

Subjects were divided randomly into groups of three. Every subject received an endowment of 60 Norwegian krone (NOK) [about \$8 US (USD)]. Their task was to decide how much of this endowment to keep for themselves and how much to contribute to a group account. Any amount contributed to the group account was doubled by the experimenters and was shared equally among the three group members.

We varied four aspects of the game: first, whether subjects could choose between different group types; second, whether the contribution decision was made conditional on others' behavior; third, how many times the game was repeated (if at all); and fourth, if the game was repeated, whether group composition was kept fixed throughout.

When subjects were faced with a choice between group types, there were two options: the red and blue groups (in the instructions, group types were called "X" and "Z", to avoid any unintended connotations of group names). The only difference between these groups was that in blue groups, each member received an extra payment of 50 NOK for themselves in addition to the payment in the public good game as explained above. Members of red groups received no such extra payment; however, for each member of a red group, a fixed amount of 50 NOK was donated to the Norwegian Red Cross.

In study 1, subjects chose their contribution levels by responding to the following question: "How many NOK will you contribute to the group account?" We will refer to this procedure as the "Regular game." In parts of studies 2–4, subjects were given the opportunity to condition their contributions on the average contributions of others in their group. As in ref. 8, subjects were first faced with an unconditional question, similar to the one above: "How many NOK will you contribute to the group account?" Next, they were asked to fill in a table representing a conditional contribution decision. For each possible average contribution from the others in the group, rounded to the nearest 5 NOK, the subject was asked to report how many NOK she herself wanted to contribute to the group account. For two randomly drawn subjects within each group, actual contributions were determined by the unconditional question; the third group member's contribution was then determined by her response to the conditional question, based on the average (unconditional) contributions of the two others. Since asking subjects such conditional questions is often called "the strategy method" (25), we refer to this procedure as the "Strategy game."

When the game was repeated within a part, subjects were told that their actual payment, as well as the donation to the Red

Cross, would be determined by the average of all the rounds in that part.

Each study consists of four parts, except study 1, which has no part II (its parts are renumbered here as compared with ref. 22 to correspond to studies 2–4). Table 1 summarizes the features used in each part of each study.

Part I was always the Regular game played once, without any informational feedback. Part II was the Strategy game played once, again without feedback. In part III, subjects first made a choice between red and blue groups before playing 10 rounds of the Regular game (studies 1, 3 and 4) or the Strategy game (study 2), with the group composition kept fixed. After each round in part III, information was provided about average contributions and earnings in the subject's own group.

In part IV, the design varied among studies. In study 1 and 4, there were 20 rounds, in each of which subjects chose between red and blue groups before playing the Regular game. Groups were rematched between rounds. After each round, information was provided about average contributions and earnings in the subject's own group, average earnings in one red group, and average earnings in one blue group. In study 2, part IV consisted of five rounds in each of which subjects chose between red and blue groups before playing the Strategy game. In study 3, part IV was identical to part II.

Results

Red Groups Are Better at Sustaining Cooperation. Since we aim to explore why red groups sustain cooperation, the first question is whether this result is replicated in studies 2–4.

As in study 1, a substantial share of subjects chose red groups in the present experiments. In part III of studies 2–4, 35% chose red groups; in part IV of study 4, an average of 31% chose red groups (Table 2 and see *SI Appendix, Table S1* for further details).

The relative popularity of red groups could hardly be explained by a hope that higher cooperation levels would outweigh the loss of the 50 extra NOK earned by members of blue groups (on average, individual payoffs were always considerably higher in blue groups in all four studies) (*SI Appendix, Fig. S2*).

Fig. 1 shows contribution levels by red and blue groups for all four studies. Note that the dots on the left illustrate average contributions in parts I and II; here, no group choice had been made, so we classify subjects as "red" or "blue" based on their later choice of group type in part III. Periods 1–10 are part III, and periods 11–30 illustrate part IV (varying in length among studies).

Fig. 1 shows that for each study the contribution levels are always higher in red groups than in blue groups; for the pooled data, the difference is strongly significant (see *SI Appendix, Fig. S3* for statistical tests accounting for multiple hypotheses).

If one disregards possible end-game effects in the last period, there are no statistically significant downward trends among red groups in part III; in blue groups, however, contributions decline substantially over time except in study 2 (see *SI Appendix, Table S3* for formal tests). In part IV of studies 1 and 4, where new groups are formed in each of the 20 rounds, a downward trend is present for both group types; however, the decline is significantly

Table 1. Design features

Study	Part I	Part II	Part III	Part IV
1	1× Regular		Group choice + (10× Regular)	20× (group choice + Regular)
2	1× Regular	1× Strategy	Group choice + (10× Strategy)	5× (group choice + Strategy)
3	1× Regular	1× Strategy	Group choice + (10× Regular)	1× Strategy
4	1× Regular	1× Strategy	Group choice + (10× Regular)	20× (group choice + Regular)

Regular: unconditional contribution decisions; Strategy: unconditional contribution decision for two subjects per group, conditional contribution decision for the third subject. Group choice: choice of red or blue group and new matching of groups.

Table 2. Group choices and profile categories

Profile	% subjects	% red
Free-rider	13	17*
Altruist	4	75*
Conditional cooperation	77	35
Other	7	43
All profiles	100	35

The table shows the distribution of conditional contribution profiles as well as part III group choices for studies 2, 3, and 4. All numbers are in percentages.

*Denotes group compositions that are different from the other groups at a 1% level of confidence using a Fischer test ($n = 324$).

steeper in blue groups than in red groups (*SI Appendix, Table S4*). Study 3 has only one period in part IV, allowing no trends.

Study 2 stands out in that there is no statistically significant upward or downward trend in contributions for either blue groups or red groups. However, in the Strategy game, which was used in parts III and IV of study 2, the incentives for unconditional contributions differ somewhat from those in the Regular game. If subject A's unconditional contribution is drawn to count in the Strategy game, another group member B's conditional contribution in the same round will be a direct function

of A's unconditional contribution. If A expects a large share of others to be conditional cooperators, she thus has an additional strategic incentive to contribute as compared with the Regular game. This must be borne in mind when interpreting results from study 2, parts III and IV.

Overall, we replicate the main findings from study 1: A substantial share of subjects choose red groups; contributions are higher in red groups than in blue groups; and contributions display a significantly steeper decline over time in blue groups than in red groups (except in study 2). We now turn to possible explanations for this pattern.

Altruists Prefer Red Groups, and Free-Riders Prefer Blue Groups. To describe subjects' conditional contribution preferences, we use the conditional contribution profiles elicited in part II.

Following the classification rules used in ref. 23, we classify subjects as altruists, free-riders, conditional cooperators, and others. Subjects who contributed less than 50% irrespective of others' contributions are classified as free-riders. Altruists are those who contributed more than 50% regardless of others' average contributions. Subjects whose conditional contribution profiles are nondeclining and cross the 50% mark are classified as "conditional cooperators," and the rest are classified as "others."

The results show a significant selection effect in the pooled data for part III (Table 2 and *SI Appendix, Table S5*). Thirteen

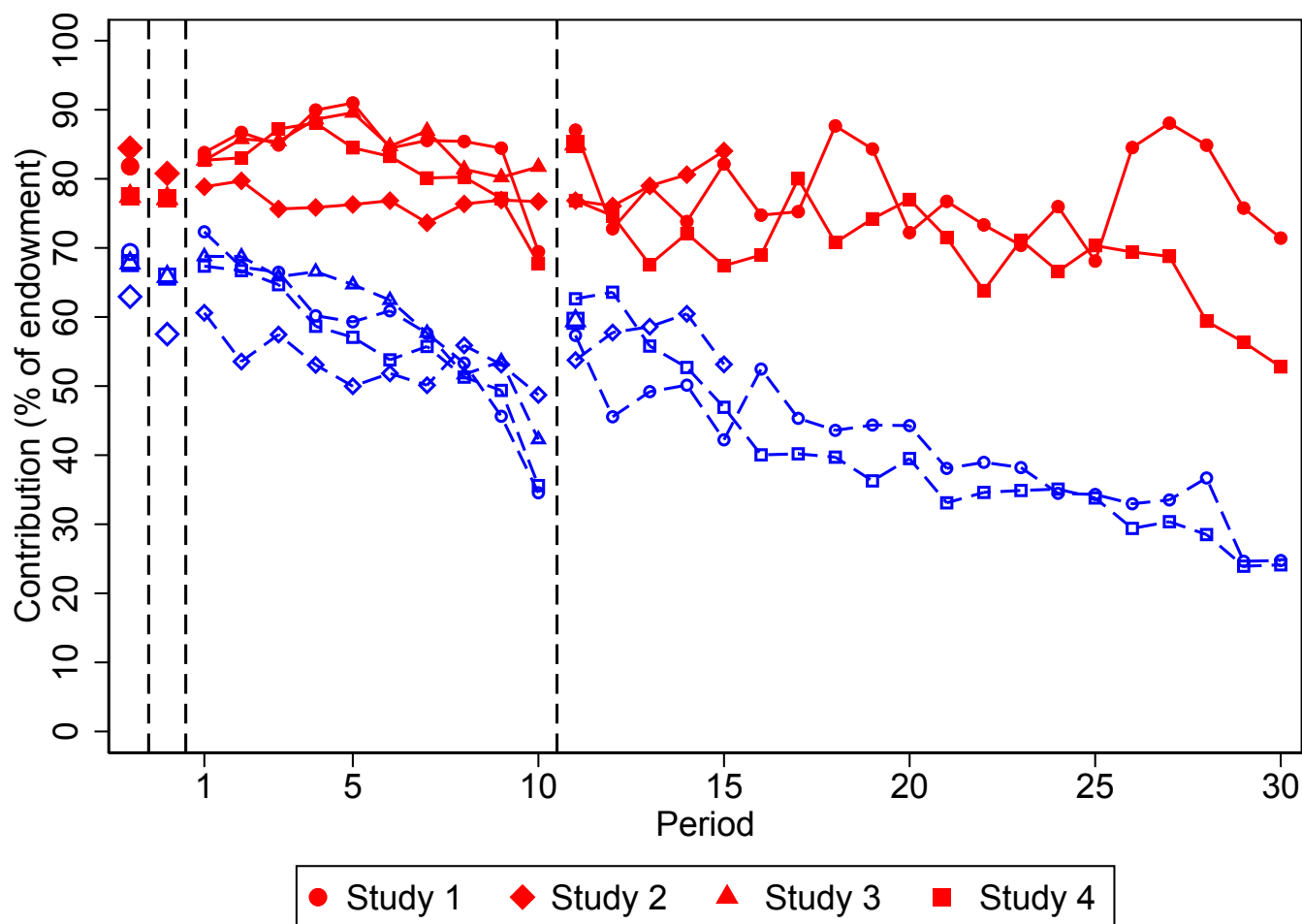


Fig. 1. Average contributions by period and group type. The graph shows average contributions in the four studies by period and group type ($n = 411$). Red groups are shown by solid lines and filled symbols; blue groups are shown by dashed lines and hollow symbols. For study 2, parts II–IV and study 3, part II, unconditional contributions from the Regular game are used. The four sections show parts I–IV. Part IV had a varying number of periods across studies. Exact values and SEs are provided in *SI Appendix*.

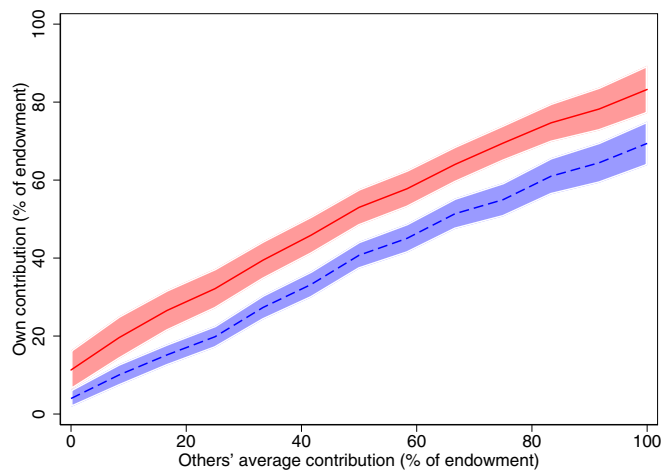


Fig. 2. Contribution profile by group type. The figure shows the average conditional contribution profiles from the Strategy game in part II and the choice of group type from part III as well as 95% confidence bands (shaded areas). The reported individual contributions levels are shown on the y axis against hypothetical contribution levels of the other group members on the x axis. Data from studies 2, 3, and 4 are included ($n = 324$).

percent of our subjects were free-riders, of whom a large majority chose blue groups. While only 4% of all subjects were classified as altruists, most of them chose red groups. The bulk of our subjects, 77%, were classified as conditional cooperators; these divided themselves between both group types, representing the majority in red groups as well as in blue groups.

As a robustness check, we also have adopted the procedure used in ref. 8, classifying subjects as free-riders, conditional cooperators, hump-shaped, and others, using these authors' classification rules. Using data from study 3 and 4, we find that free-riders are significantly less likely to choose red groups (*SI Appendix, Tables S6 and S7*).

If a larger share of altruists stimulate conditional cooperators' contributions in red groups, while a larger share of free-riders undermine conditional cooperators' contributions in blue groups, we should expect that contributions in red and blue groups diverge over time even if we include only contributions from conditional cooperators. This is indeed what we observe (*SI Appendix, Fig. S4*). Notice, however, that even in part I the conditional cooperators in red groups start off contributing more than the conditional cooperators in blue groups.

Levels Are Higher in Red Groups, but Slopes Are Similar in Red and Blue Groups. The classification used above may be too broad to capture all interesting aspects of self-selection. Below, we instead describe subjects' conditional contribution profiles by approximating each subject's conditional contribution profile linearly and then computing the slope and average contribution level of each subject's conditional contribution profile.

It turns out that the slope is essentially unrelated to self-selection into red groups in part III. In the pooled data, we find no significant differences between the slopes of the conditional contribution profiles in the red and blue groups. However, subjects with higher average contribution levels are much more likely to choose red groups (for details, see *SI Appendix, Table S8*).

Fig. 2 presents the average conditional contribution profiles in red and blue groups. For every possible level of others' contributions, contributions in red groups lie strictly above the 95% confidence band for contributions in blue groups. The slopes, on the other hand, appear rather similar.

Fig. 3 illustrates the distribution of average conditional contribution levels in red and blue groups (measured as percentages

of all members in red and blue groups, respectively). While the average conditional contribution levels vary widely within each group type, the share of subjects with low levels is higher in blue groups, while the share of subjects with high levels is higher in red groups.

Alternative Hypotheses. The results so far are consistent with the explanation that more of the generous individuals self-select into red groups, which triggers different dynamic responses from conditional cooperators in red and blue groups. However, other explanations may be possible.

First, individual preferences might change over time. For example, a person classified as an altruist in part II may find, through disappointment with others' contributions during part III, that her preferences are more conditional than she initially felt. Using data from study 3, in which parts II and IV are identical, we can compare subjects' conditional contribution profiles before and after part III. As shown in Table 3, the classification of subjects is quite stable: Only 9 of 90 subjects changed their profile classification between part II and part IV.

Second, the self-selection pattern may change over time, so that red groups gradually attract more cooperative types. This is not the case, however: In study 4, part IV, neither the slope nor the average level of conditional contribution profiles changed significantly over time for either group type (see details in *SI Appendix, Fig. S5*).

Finally, since we have focused on part III, where fixed groups play a repeated game, strategic reasons may interact with the group and/or contribution choices in ways that help red groups sustain their cooperation. To consider this, we can look at part IV of studies 4 and 1, in which groups are rematched between each round.

Fig. 1 indicates that cooperation in red groups is still relatively stable in part IV of studies 4 and 1. Statistical analysis (*SI Appendix, Table S4*), however, reveals a statistically significant negative trend even in red groups. Nevertheless, the negative trend is significantly stronger in blue groups, meaning that strategic concerns alone can hardly explain red groups' superior ability to keep cooperating over time.

Discussion

If altruists help maintain the cooperation of the red groups, while free-riders undermine it in blue groups, we would expect to see that, regardless of whether the groups are red or blue, the

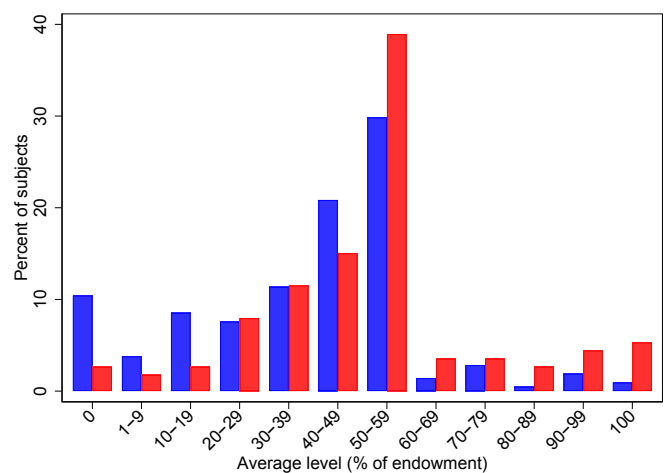


Fig. 3. Average conditional contributions in the red and blue groups. The graph shows average conditional contribution levels by group choice (average levels of 0, 100, and 10-unit intervals between 0 and 100) as the percentage of total endowment. Average conditional contributions levels are calculated as the mean level of the conditional contribution profile reported in part II. Group choice is from part III. Data are from studies 2, 3, and 4 ($n = 324$).

Table 3. Transitions between profile categories

Part II profile	Part IV profile				Total
	Free-rider	Altruist	Conditional cooperation	Other	
Free-rider	74	0	19	7	100
Altruist	0	80	20	0	100
Conditional cooperation	3	3	88	6	100
Other	20	0	27	53	100
All profiles	13	6	71	10	100

The table shows the transitions between profile categories from part II to part IV in study 2 and 3. In study 2, we use period 5 of part IV. All numbers are in percentages.

groups with altruist members are generally better at sustaining cooperation than groups with free-riders. This is indeed what we observe (*SI Appendix, Fig. S8*).

If a group's commitment to a global public good (here, the Red Cross) attracts more cooperative group members, this may have important implications outside the laboratory, e.g., for firms' social responsibility. In teamwork contexts where employers cannot observe workers' extra efforts, such efforts essentially represent voluntary contributions to the team's production. If firms with a reputation for being socially responsible attract more cooperative workers (26), these workers may have profound effects on firm productivity through their supportive effect on other workers' good will, thus stimulating a more productive firm culture. Previous studies have found that in teamwork contexts without individual performance pay, corporate social responsibility can attract more cooperative workers and increase productivity (27–31).

We did, in fact, ask our subjects about their views on corporate social responsibility in a postexperimental survey. The results (*SI Appendix, Table S9*) show that subjects who reported being more strongly concerned about firms' social responsibility in their future job choices than with the other job characteristics included in the survey (e.g., pay; predictability; ability to influence one's own workdays) were significantly and substantially more likely to choose red groups in the experiments.

By a similar logic, organizations and clubs might, through costly commitments to global public goods such as biodiversity or climate change mitigation, be able to attract more cooperative members whose presence would increase other members' willingness to cooperate.

Recent reports demonstrate that the option to avoid giving situations can reduce overall giving (e.g., refs. 32 and 33). Although the context of our experiments is different, our results complement this strand of research by showing that self-sorting may be associated with cooperativeness along several dimensions.

Conclusions

Groups precommitted to a charity attract more cooperative group members who are better able to sustain the group's cooperation over time. Our present experiments confirm this previous finding, exploring the underlying mechanisms at hand.

Group dynamics are likely to be strongly influenced by conditional cooperators, who constitute a large majority of our subjects.

We find that those who self-select into the charity groups have a higher average conditional contribution level, while the slope of their conditional contribution profiles does not differ significantly from that of other subjects. Our results indicate that charity groups are better able to sustain their cooperation because they attract more generous members who, in turn, have a positive influence on their conditionally cooperative peers.

Outside the laboratory, our results contribute to the understanding of corporate social responsibility (27, 30): When individual teamwork efforts are unobservable, extra efforts are essentially voluntary contributions to the team's production. If the mechanisms at work are similar to those studied in our experiment, the higher productivity of corporate socially responsible firms may be the result of such firms' ability to attract more altruistic workers combined with these workers' stimulating effect on the cooperativeness of their conditionally cooperative coworkers.

Materials and Methods

The collected data were anonymous and cannot be linked to individual (natural) persons. Procedures required for such data collection by The Personal Data Act in Norway have been followed. Recruitment procedures to the experiments included information that participation was voluntary. Subjects were recruited at the University of Oslo, and in some of the experiments the subjects enrolled using the recruitment program Expmotor, which was developed by Erik Sørensen and Trond Halvorsen at the Norwegian School of Economics. All experiments were programmed in z-tree (34) and were conducted at the Oeconlab, the experimental laboratory of the Department of Economics at the University of Oslo. All analyses were conducted in Stata version 15. The data and the estimation code are available in *SI Appendix*.

All procedures were fully anonymous. At the beginning of each part of the experiment, instructions including all the rules for that part were handed out and were read aloud (*SI Appendix*), and participants answered control questions to ensure their understanding of the instructions. Payments were handled by people who did not know the details or the objectives of the experiments or any choices made by the subjects during the experiment.

Groups were formed randomly. When subjects were given a choice of group type, groups were formed randomly among those who chose red and blue groups, respectively. If the number of subjects preferring a group type was not divisible by three, one group was mixed, and its type was determined by the majority preference. Subjects were informed whether their group was mixed. Contribution data from mixed groups are excluded from the analysis.

Study 1 was conducted in February 2008, comprising five sessions with 87 subjects. A more detailed description of the design and results can be found in ref. 22. Sessions lasted about 1 h, yielding an average payment per subject of roughly 450 NOK (equivalent to 83.50 USD at the time of the experiment). Study 2 was conducted in March 2017. There were five sessions of ~90 min each with 117 participants in total. The average payment per subject was 455 NOK (~53.50 USD at the time of the experiment). Study 3 was conducted in October 2017. There were four sessions with a total of 90 participants, lasting 65–70 min on average. The average payment per subject was 438 NOK (~55 USD at the time of the experiment). Study 4 was conducted in April 2018, comprising five sessions with a total of 117 participants, lasting 74 min on average and yielding an average payment per subject of 454 NOK.

ACKNOWLEDGMENTS. We thank two anonymous referees for thorough and valuable comments and Johannes Alsvik, Tora H. Knutsen, Christian Presterud, and Ingvild Ruen for valuable research assistance. K.E.H. received funding from Research Council of Norway Grant 268 174 and internal funding from the Frisch Centre. This work was supported by the Fund for Applied Research, by internal funds from the Department of Economics, University of Oslo, by the Oslo Centre for Research on Environmentally Friendly Energy (CREE), and by Research Council of Norway Grant 209698. K.A.B., K.E.H., and K.N. are members of CREE.

- Barrett S (2003) *Environment and Statecraft: The Strategy of Environmental Treaty-Making: The Strategy of Environmental Treaty-making* (Oxford Univ Press, Oxford).
- Holmstrom B (1982) Moral hazard in teams. *Bell J Econ* 13:324–340.
- Brekke KA, Nyborg K, Rege M (2007) The fear of exclusion: Individual effort when group formation is endogenous. *Scand J Econ* 109:531–550.
- Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge Univ Press, Cambridge, UK).
- Chaudhuri A (2011) Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Exp Econ* 14:47–83.
- Ledyard JO (1995) Public goods: A survey of experimental research. *The Handbook of Experimental Economics*, eds Kagel JH, Roth AE (Princeton Univ Press, Princeton), pp 111–194.
- Falk A, Fischbacher URS, Gächter S (2013) Living in two neighborhoods—Social interaction in the lab. *Econ Inq* 51:563–578.
- Fischbacher U, Gächter S, Fehr E (2001) Are people conditionally cooperative? Evidence from a public goods experiment. *Econ Lett* 71:397–404.
- Frey B, Meier S (2004) Social comparisons and pro-social behavior: Testing “conditional cooperation” in a field experiment. *Am Econ Rev* 94:1717–1722.

