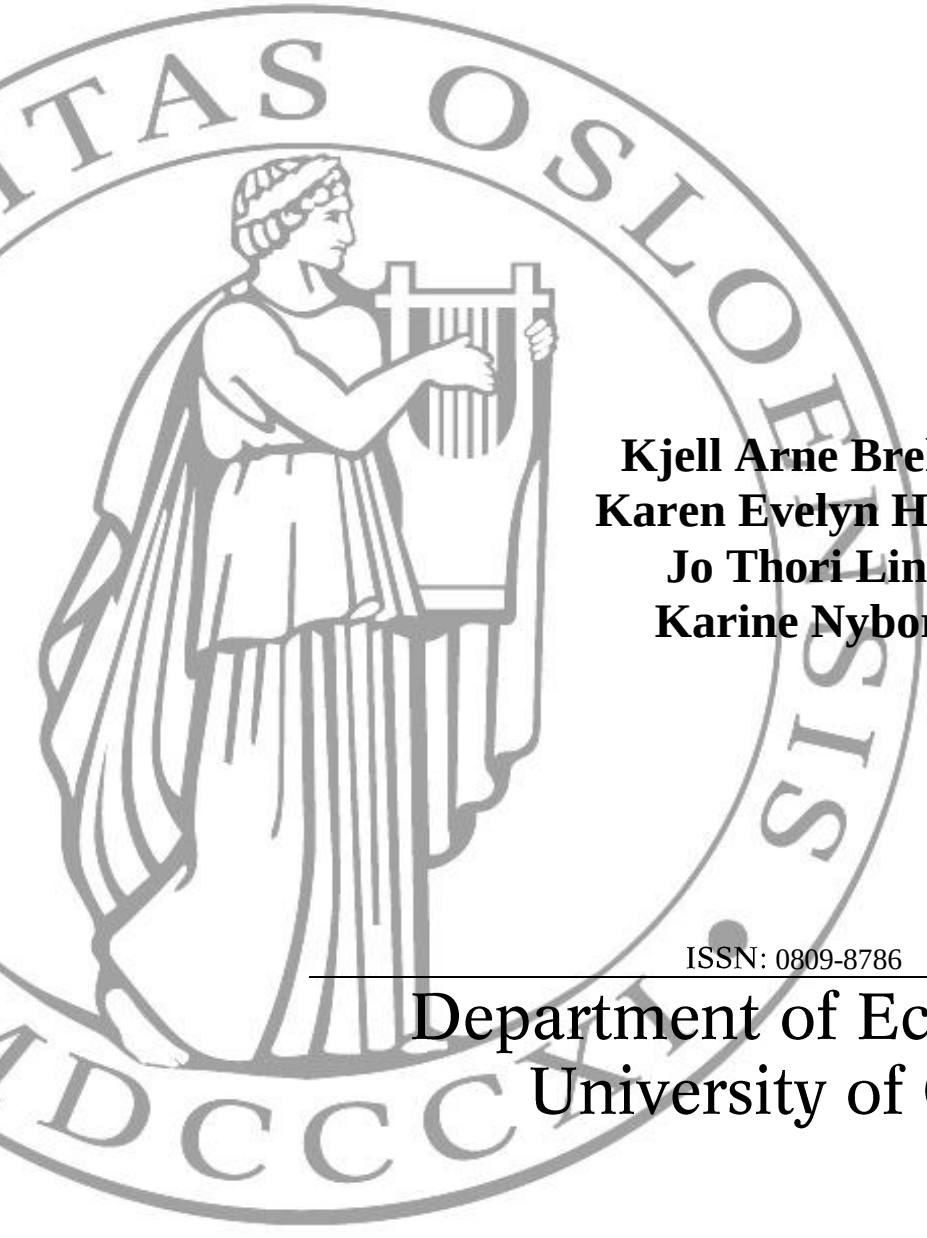
MEMORANDUM

No 08/2009

Playing with the Good Guys: A Public Good Game with Endogenous Group Formation



Kjell Arne Brekke
Karen Evelyn Hauge
Jo Thori Lind
Karine Nyborg

ISSN: 0809-8786

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.oekonomi.uio.no
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no
e-mail: frisch@frisch.uio.no

Last 10 Memoranda

No 07/09	Benedicte Carlsen and Karine Nyborg *The Gate is Open: Primary Care Physicians as Social Security Gatekeepers*
No 06/09	Alessandro Corsi and Steinar Strøm *The Premium for Organic Wines? Estimating a Hedonic Price Equation from the Producer Side*
No 05/09	Jo Thori Lind, Karl Moene and Fredrik Willumsen *Opium for the Masses? Conflict-induced Narcotics Production in Afghanistan*
No 04/09	Jo Thori Lind and Karl Moene *Misrely Developments*
No 03/09	Steinar Holden and Fredrik Wulfsberg *Wage Rigidity, Institutions, and Inflation*
No 02/09	Nils-Henrik M. von der Fehr and Petter Vegard Hansen *Electricity Retailing in Norway*
No 01/09	Diderik Lund *Rent Taxation for Nonrenewable Resources*
No 30/08	Finn R. Førsund *Good Modelling of Bad Outputs: Pollution and Multiple-Output Production*
No 29/08	Michael Hoel *Bush Meets Hotelling: Effects of Improved Renewable Energy Technology on Greenhouse Gas Emissions*
No 28/08	Eric Nævdal and Jon Vislie *Climate Change, Catastrophic Risk and the Relative Unimportance of Discounting*

Previous issues of the memo-series are available in a PDF® format at:
http://www.oekonomi.uio.no/memo/index.html

Playing with the Good Guys

A Public Good Game with Endogenous Group Formation*

Kjell Arne Brekke, Karen Evelyn Hauge, Jo Thori Lind, Karine Nyborg †
Department of Economics, University of Oslo

April 17, 2009

Abstract

In public good games, voluntary contributions tend to start off high and decline as the game is repeated. If high contributors are matched, however, contributions tend to stay high. We propose a formalization predicting that high contributors will self-select into groups committed to charitable giving. Testing this experimentally, we let subjects choose between two group types, where one type donate a fixed amount to a charity. Contributions in these groups stayed high, whereas contributions in the other groups showed the well known declining pattern. One implication is that corporate social responsibility may attract more responsible employees.

JEL codes: D11, D12, D64, H41
Keywords: Altruism; conditional cooperation; self-selection

1 Introduction

Charitable giving is much more widespread in real life than one would expect assuming purely selfish behavior (see e.g. Andreoni, 1988a, 1990). This willingness to contribute voluntarily is confirmed by substantial evidence from public good game experiments (e.g. Ledyard, 1995; Fehr and Gächter, 2000, 2002). A striking finding from such experiments,

*We are grateful to the Research Council of Norway (RCS) for funding through the RAMBU/Miljø2015 programmes, to Geir Asheim and numerous conference and seminar participants for comments and suggestions, and to Kristine Korneliussen for excellent research assistance. Part of this project was undertaken while Brekke and Nyborg were employed by the Ragnar Frisch Centre for Economic Research. The authors are part of the ESOP - Centre of Equality, Social Organization, and Performance, which is funded by the RCS as a Norwegian Centre of Excellence.

†PB 1095 Blindern, 0317 Oslo, Norway. Emails: k.a.brekke@econ.uio.no, k.e.hauge@econ.uio.no, j.t.lind@econ.uio.no, karine.nyborg@econ.uio.no.

however, is that although average contributions usually start off at high levels, they tend to decline dramatically as games are repeated (ibid.). This decline in contributions is found both when group compositions are held fixed (partner designs) and change (stranger design) over time (Andreoni, 1988b; Fehr and Gächter, 2000; Keser and van Winden, 2000; Croson, 1996).

The ability to sustain voluntary contributions is important in many economic contexts. In the case of teamwork with unobservable individual effort (Holmström, 1982), for instance, individuals' work effort can be regarded as voluntary contributions, and maintaining a high level of cooperativeness in such teams can be crucial to firms' survival. A better understanding of how and why voluntary contributions can (or cannot) be sustained is thus of general interest.

In the present paper, we contribute to this goal by first outlining a theoretical model of contributions in a public game context, based on the idea that individuals strive towards some conception of a morally ideal contribution, but that these ideals are affected by others' behavior. Within this framework, we show that contributions are likely to decline over time, unless high contributors are matched with each other. Moreover, we show that if individuals' willingness to contribute are correlated across different public goods, high contributors will self-select into groups for which a pre-commitment to charitable giving is announced. We test our predictions by means of a public good game experiment where subjects can choose which type of group to join: either a group where a fixed amount is given to the Red Cross, or a group where the same amount is given to subjects themselves. We find that high contributions are sustained in the former group type, while for the latter, the usual declining pattern is observed. Hence our empirical findings support the theoretical predictions.

To our knowledge, few attempts have been made to explain, in the language of formal theoretical models, the initially high, but declining contributions pattern in repeated public good games. Nevertheless, previous experimental findings indicate that this phenomenon may be largely due to conditional cooperators; that is, subjects who start off contributing relatively high amounts, but who condition their further cooperation on others' behavior. In Fischbacher and Gächter (2006), for example, 55 percent of the subjects were classified as conditional cooperators, that is, subjects who contribute more the more others contribute; 23 percent were classified as free-riders, while less than 2 percent were unconditional cooperators. Moreover, the authors find that the heterogeneity of types explains a large part of the decay in contributions. Similarly, Fischbacher, Gächter, and Fehr (2001) find that 50 percent of their subjects are conditional cooperators, while 30 percent are free-riders. Hauge (2007) asked her subjects about what they believed to be the morally ideal contribution; the majority of her subjects reported a conditional moral ideal, approximately equal to the average contribution of the others in their group. Neugebauer, Perote, Schmidt, and Loos (2009), using an experimental set-up designed to distinguish conditional cooperation from

strategic motives as well as the 'false consensus effect', conclude that the decline in cooperation appears to be driven by selfish-biased conditional cooperation and adaptive belief learning. Fischbacher and Gächter (forthcoming) find that on average, people are "imperfect" conditional cooperators: contributions depend both on subjects' preferences and their beliefs about others' contributions, but subjects do not fully match others' expected contributions. In line with these findings, the theoretical framework presented below assumes that individuals strive towards their own moral ideals, but that these ideals are conditioned on others' behavior.

Our theoretical model predicts that if high contributors are matched with each other, high average contribution levels can be sustained. This result has been confirmed by several previous experiments where, without subjects' knowledge, groups are formed exogenously based on previous contribution behavior (Gächter and Thöni, 2005; Gunnthorsdottir, Houser, and McCabe, 2007; Ones and Putterman, 2007). While these studies use exogenous group formation, the most common case in everyday life is that people self-select into groups; nevertheless, few scholars have studied endogenous group formation within the context of public good games.[1] If some groups succeed in sustaining cooperation, free-riders will have an incentive to invade those groups. This is, indeed, reflected in the endogenous group formation game investigated by Ehrhart and Keser (1999), who observe that "the more cooperative subjects are continually on the run from the less cooperative ones. However, the less cooperative ones keep following them around" (op.cit, p.1). One route to sustain cooperation is to limit access to (or exit from) endogenously formed groups (Ahn, Isaac, and Salmon, 2008; Page, Putterman, and Unel, 2005).

In the present paper, we suggest another way for high contributors to match. Our proposal is partly based on Brekke and Nyborg (2008), who demonstrate that if cooperative behavior is derived from an underlying ethical principle, while the weight attached to this principle varies between individuals, one would expect to find a positive correlation between individuals' cooperativeness in different contexts. Exploring this idea and assuming that individual preferences to contribute to different public goods are positively correlated, we show theoretically that high contributors will self-select into groups pre-committed to charity contributions, and that these groups will be able to sustain cooperation.

We test this prediction empirically by means of a public good game with endogenous group formation, where subjects self-select into groups with different payment structures. The experiment consists of three parts. The first part is a standard one-shot public good game experiment with exogenous group formation. In the second part, subjects can choose between two group types, red and blue. The only difference is that in red groups, a fixed amount of money is donated to the Red Cross; while in blue groups, the same amount is,

[1] One exception is Gürerk and Rockenbach (2006), who focus on subjects' choice between sanctioning and non-sanctioning institutions. In their experiment there are only two groups, and group size is endogenous.

instead, given to individual group members. In Part 2 the game is repeated ten times, and the group composition remains fixed for all 10 periods. In Part 3, subjects decide on their preferred group type between each period, and new groups are formed in every period; moreover, the number of periods is increased to 20.

We find that throughout Parts 2 and 3, a substantial share of subjects, about forty percent, choose to be in red groups. Contributions in red groups are, on average, substantially higher than in blue groups. However, the difference is insufficient to leave red group membership more profitable in monetary terms, implying that red group membership can hardly be explained by expectations of higher pecuniary payoff. Interestingly, while the blue groups display the usual deterioration of contributions over time, no clear downwards trend is observed in the red groups.

Our results may be relevant for several real life phenomena, of which one, returning to the case of teamwork, is corporate social responsibility. A positive correlation between individuals' preference to contribute to local and global public goods would imply that workers who are less likely to shirk in teamwork are also more willing to accept lower wages in return for a socially responsible employer. Brekke and Nyborg (2008) show that this may enable socially responsible firms not only to survive, but possibly even to drive non-responsible firms out of business altogether – due to their ability to attract more responsible, and hence more hard-working, employees. Our experimental results support the empirical relevance of this idea.

2 Theory and predictions

Consider a public good game with n members in each group, where each player is given an exogenous initial endowment Y. The game is repeated T times and that group composition is fixed throughout T. Let t denote period number ($1 \leq t \leq T$). Each player decides simultaneously and anonymously how much of Y to contribute to her group in period t, g_i^t, keeping the remaining part of Y for herself. Contributions to the group are multiplied by a factor M ($1 < M < n$), and then shared equally between group members. Consequently, g_i^t is a voluntary contribution to a local public good.

2.1 Choice of contribution

Let each individual i have preferences of the following kind:

$$u_i^t = x_i^t + \beta_i g_i^t - \frac{\alpha}{2}(g_i^t - \hat{g}_i^t)^2 \tag{1}$$

where $\alpha > 0$. Here, u_i^t is individual i's utility in period t, x_i^t denotes i's monetary payoff in period t, while the other terms reflect social preferences. β_i is the unconditional warm glow of giving (Andreoni, 1990) experienced by i per unit contributed to the local public good. The last term can be considered a *conditional* warm glow (Konow, 2009), which depends not only on i's own contribution g_i^t, but also on the level of some yardstick $\hat{g}_i^t$. This is similar to the self-image model of Brekke, Nyborg, and Kverndokk (2003), who interpreted $\hat{g}_i^t$ as i's idea about the morally ideal contribution, which they assumed, in turn, to be determined through a Kantian-style ethical argument. Although our assumptions about the determination of $\hat{g}_i^t$ will be different, the interpretation is similar: we consider $\hat{g}_i^t$ to be i's idea of the morally ideal contribution, implying that conditional warm glow is higher the closer i's actual contribution is to her perception of what she should, ideally, have contributed. Note that, in accordance with the social psychological theory of cognitive dissonance (Festinger, 1957), conditional warm glow is always negative, but less so when g_i^t is closer to $\hat{g}_i^t$.

In period t, monetary payoff to i is given by

$$x_i^t = Y - g_i^t + \frac{M}{n}\sum_{j=1}^{n} g_j^t \tag{2}$$

where $j = 1, ..., n$ are the members of i's group. This implies (rearranging and inserting into (1)) that utility can be written as

$$u_i^t = Y + \frac{M}{n}\sum_{j\neq i} g_j^t + \left(\beta_i - \frac{n-M}{n}\right) g_i^t - \frac{\alpha}{2}(g_i^t - \hat{g}_i^t)^2 \tag{3}$$

Assume, first, that $\hat{g}_i^t$ and g_j^t $(j \neq i)$ are exogenous. Since the finite time horizon rules out Folk Theorem-type equilibria, the utility maximizing contribution can be solved separately for each period. The first order condition for an interior solution requires that marginal warm glow equals the net monetary cost of a marginal contribution:

$$\beta_i - \alpha\left(g_i^t - \hat{g}_i^t\right) = \frac{n-M}{n}. \tag{4}$$

If the solution is not interior, we must take into account that g_i^t can neither be negative nor exceed the initial endowment Y. The utility-maximizing contribution can thus be written as follows:

$$g_i^t = \begin{cases} 0 & \text{if} \quad \hat{g}_i^t + \frac{1}{\alpha}\left(\beta_i - \frac{n-M}{n}\right) < 0 \\ \hat{g}_i^t + \frac{1}{\alpha}\left(\beta_i - \frac{n-M}{n}\right) & \text{if} \quad 0 \leq \hat{g}_i^t + \frac{1}{\alpha}\left(\beta_i - \frac{n-M}{n}\right) \leq Y \\ Y & \text{if} \quad Y < \hat{g}_i^t + \frac{1}{\alpha}\left(\beta_i - \frac{n-M}{n}\right) \end{cases} \tag{5}$$

For interior solutions, the optimal contribution is strictly increasing in the ideal $\hat{g}_i^t$ and in β_i.

To improve the readability of the results presented below, we will focus on interior solutions. While corner solutions cannot, of course, be precluded in general, all our main results in what follows can be generalized to allow for corner solutions simply by replacing the parameter β_i by a truncated parameter $\tilde{\beta}_i(\hat{g}_i^t)$, defined as follows:

$$\tilde{\beta}_i(\hat{g}^t) = \begin{cases} \frac{n-M}{n} - \alpha\hat{g}_i^t & \text{if} \quad \beta_i < \frac{n-M}{n} - \alpha\hat{g}_i^t \\ \beta_i & \text{if} \quad \frac{n-M}{n} - \alpha\hat{g}_i^t < \beta_i < \frac{n-M}{n} + \alpha(Y - \hat{g}_i^t) \\ \frac{n-M}{n} + \alpha(Y - \hat{g}_i^t) & \text{if} \quad \frac{n-M}{n} + \alpha(Y - \hat{g}_i^t) < \beta_i \end{cases} \tag{6}$$

Replacing β_i by $\tilde{\beta}_i(\hat{g}_i^t)$ in the middle line of eq. (5) would yield

$$g_i^t = \hat{g}_i^t + \frac{1}{\alpha}\left(\tilde{\beta}_i(\hat{g}_i^t) - \frac{n-M}{n}\right) \tag{7}$$

which gives the optimal g_i^t for every individual, even in cases of corner solutions. Other results provided below can be generalized similarly.

2.2 Conditional ideals

Let us now consider the idea that the morally ideal contribution $\hat{g}_i^t$ may depend on others' contributions. As mentioned above, there is considerable experimental evidence for such a relationship (Fischbacher and Gächter, 2006; Fischbacher, Gächter, and Fehr, 2001; Fischbacher and Gächter, forthcoming); in particular, Hauge (2007), asking subjects in a public good game experiment directly, found that a majority of subjects reported moral ideals approximately equal to the average of others' contributions.

Suppose now that $\hat{g}_i^{t+1}$, i's idea of her morally ideal contribution in period $t+1$, is a weighted average of her previous ideal $\hat{g}_i^t$ and the mean contribution in her group in the previous period, $\bar{g}_i^t$. That is, if $\lambda \in [0,1]$ is the weight placed on the previous ideal,

$$\hat{g}_i^{t+1} = (1-\lambda)\bar{g}_i^t + \lambda\hat{g}_i^t. \tag{8}$$

For $t > 1$, the ideal $\hat{g}_i^t$ and thus also the marginal warm glow from contributions to the local public good will now be increasing in others' contributions, and i will tend to be what Fischbacher, Gächter, and Fehr (2001) classify as a conditional cooperator.

If individuals are forward-looking, eq. (8) will give rise to strategic considerations: person j may want to contribute high amounts in early periods in order to increase others' ideals, and thus their contributions, in later periods. While such strategic interaction complicates the analysis considerably, experimental evidence indicates that this is not the most important force at play in public good games: testing this explicitly, Neugebauer, Perote, Schmidt, and Loos (2009) rejected the hypothesis that strategic play is the driving force of the decay in

contributions over time, concluding that "[t]he only viable hypothesis according to our data is the one of conditional cooperation and adaptive belief learning" (op.cit, p.57). Thus, below we will simplify the analysis by making the following assumption; individuals are myopic in the sense that they do not take into account that ideals will change over time; more specifically, they act as if $\hat{g}_j^t = \hat{g}_j^s$ for all $s > t$ and for $j = 1, ..., n$.

From (8), we note that the ideal will be stable, increasing or decreasing in t when ${\bar{g}_i}^t = \hat{g}_i^t$, $\bar{g}_i^t > \hat{g}_i^t$, or $\bar{g}_i^t < \hat{g}_i^t$ respectively. The latter is equivalent to (due to eq. (4))

$$\bar{\beta}_i - \frac{n-M}{n} < 0 \tag{9}$$

where $\bar{\beta}_i$ is the mean of β_j in i's group.

Proposition 1. *Assume that $\hat{g}_i^t$ changes according to Equation (8) and that the utility-maximizing contribution is interior for all i. Then, individual and average contributions will decline from t to $t+1$ whenever $\bar{\beta}_i < \frac{n-M}{n}$.*[2]

Given that the ideal evolves over time conditional on average group contributions, Proposition 1 states a condition for how contributions themselves will evolve over time. Specifically, when the average β_i is sufficiently low, and the initial ideal $\hat{g}_i^t$ is sufficiently high, contributions will decline over time. A high initial ideal may either be caused by unrealistically high expectations of others' contributions, or by moral principles such as the one Sugden (1984) calls "the principle of reciprocity": i has a moral obligation to contribute what she would have preferred everyone to contribute, unless others in her group contributes less than this; in the latter case, i's responsibility is limited to providing the least amount contributed by others in her group.

2.3 Choice of group type: self-selection

The above Proposition implies that if a group consists solely of individuals with $\bar{\beta}_i \geq \frac{n-M}{n}$, average contributions will stay high, and may even increase, over time. Our next topic is, thus, whether individuals with high β_i values can self-select into the same groups without being invaded by free-riders. We will show that if some groups are pre-committed to charity contributions, this may in fact be the case.

Assume that certain groups, which we will call red groups, are pre-committed to charity in the sense that a person's membership in a red group secures an exogenously fixed donation

[2] Allowing for corner solutions, the condition is

$$\bar{\beta}(\hat{g}_i^t) < \frac{n-M}{n}$$

where $\bar{\beta}(\hat{g}_i^t)$ is the mean of the truncated coefficient given in eq. (6).

D (per member per period) to a charity. Assume also that membership in any other group, called blue groups, provides the subject with an equivalent amount D to keep for herself (in addition to payoffs as given in eq. (2) above). The charity can be thought of as a global public good. In blue groups, D cannot be used for contributions to one's local public good (the group). Let $\tau_i = R$ denote that player i chooses membership in a red group, and $\tau_i = B$ otherwise.

Assume now that each individual i has preferences as follows:

$$u_i^t = \begin{cases} x_i^t + \beta_i g_i^t - \frac{\alpha}{2}(g_i^t - \hat{g}_i^t)^2 + \theta_i & \text{if} \quad \tau_i = R \\ x_i^t + \beta_i g_i^t - \frac{\alpha}{2}(g_i^t - \hat{g}_i^t)^2 & \text{if} \quad \tau_i = B \end{cases} \tag{10}$$

These preferences are exactly as above except for the term θ_i, which measures the unconditional warm glow of giving experienced by i when donating D to the global public good (the charity).[3]

Assume further that groups consist of individuals with the same group type, and as before that group compositions are constant throughout. The monetary payoff to i in period t is now given by

$$x_i^t = \begin{cases} Y - g_i^t + \frac{M}{n}\sum_{j=1}^n g_j^t & \text{if} \quad \tau_i = R \\ Y - g_i^t + \frac{M}{n}\sum_{j=1}^n g_j^t + D & \text{if} \quad \tau_i = B \end{cases} \tag{11}$$

When considering whether to choose a red or a blue group, the individual will consider the utility she gets from each alternative. This depends, of course, on the extra monetary payoff of D per period in blue groups and the warm glow per period θ_i in a red group. In addition, payoffs will depend on the behavior of others in one's group.

As a starting point, assume that i does not expect systematic differences between the behavior of subjects in red and blue groups. Assume further that her initial ideal $\hat{g}_i^1$ is exogenously given. Due to myopia, she will also disregard any changes in her own and others' ideals over time when making her group choice; hence the only difference between blue and red group membership, judged from this myopic position, is that blue groups provide an extra income D per period, while red groups provide an extra warm glow θ_i per period. Hence, i will choose red ($\tau_i = R$) if and only if $\theta_i \geq D$.

Any individual variation in group choice, and hence the global public good donation choice τ_i is then determined by the parameter θ_i. Note specifically that with the above assumptions, the optimal choice of τ_i is independent of i's own contributions to the local public good, g_i^t. Since the choices of g_i^t and τ_i are independent, the only reason to expect a

[3]There may, of course, also be a conditional warm glow associated with contributions to the *global* public good. Since that good is relevant for people inside and outside the lab alike, it seems plausible that this would be conditioned on all contributions, not only contributions from one's group or from inside the lab. Since contributions made in the experiment can change global contributions only marginally, we will assume that the "ideal contribution" to the global public good can be considered fixed, which means that both conditional and unconditional warm glow benefits will be captured by θ_i.

correlation between g_i^t and τ_i in the present model would be if β_i and θ_i were correlated.

As mentioned briefly above, there are in fact reasons to expect such a correlation. Brekke and Nyborg (2008) demonstrate that if voluntary contributions to a global (environmental quality) and a local (team production) public good originate from a common underlying ethical principle, while the strength of the concern for this principle varies between individuals, an individual's contributions to the two public goods will be positively correlated. The ethical principle they use is inspired by Kant's categorical imperative (namely, that i's moral satisfaction is increasing in the social welfare effects in the hypothetical case that everyone acted just like i); they argue, however, that the same would be true for other underlying ethical principles as well. With the simpler preference structure proposed in eq. (10), Brekke and Nyborg's argument would translate into a positive correlation between β_i and θ_i.

A positive correlation between β_i and θ_i would have several interesting implications. Below we explore some of them. Indeed, several recent experimental studies do support the idea of such a correlation (Altmann, Dohmen, and Wibral, 2008; Blanco, Engelmann, and Normann, 2007; Johnson, Dawes, Fowler, McElreath, and Smirnov, 2009). Theorem 1 is the main formal result, while the propositions point more explicitly to hypotheses to be tested empirically.

Theorem 1. *If β_i and θ_i are positive, monotone transformations of underlying stochastic variables b_i and y_i, where b_i and y_i are normally distributed and strictly positively correlated, then $\beta_i|(\theta_i \geq \hat{\theta})$ will stochastically strictly dominate $\beta_i|(\theta_i < \hat{\theta})$, for any threshold $\hat{\theta}$.*

For proof, see Appendix A.1.

In the following we will invoke the premise of the proposition as an assumption. Note that allowing β_i and θ_i to be monotone transforms of a joint normal distribution shows that the claim is true for a very large class of probability distributions.

Our first corollary shows that when β_i and θ_i are positively correlated, expected contributions[4] to the local public good, g_i^t, are higher, all else given, for those who choose red, $\tau_i = R$, than for those who choose blue, $\tau_i = B$.

Proposition 2. *For any given ideal contribution $\hat{g}_i^t$,*

$$E\left[g_i^t|\tau_i = R\right] > E\left[g_i^t|\tau_i = B\right]$$

provided that $0 < g_i^t < Y$ for some β_i in the support.

Proof. From the discussion above we know that the event $\{\tau_i = R\} \Leftrightarrow \left\{\theta_i > \hat{\theta}_i\right\}$ with $\hat{\theta}_i = D$ Thus the Proposition applies, and $\beta_i|\tau_i = R$ stochastically strictly dominates $\beta_i|\tau_i = B$. In

[4]Expectations here and in the following are to be interpreted as the expectations of an outside observer, that is, the average value if the game is played many times with different individuals drawn from the same population.

particular this dominance applies in the part of the support where g_i^t is interior. Since interior contributions are linearly increasing in β_i, the claim follows. □

2.4 Group specific conditional ideals

Let us now study what the model predicts as time evolves. Assume now that eq. (8) holds, and that individuals are myopic and do not, at the outset, expect different behavior from members of blue and red groups.[5]

From Proposition 1 above, we know that β_i, or rather its group average, is the crucial parameter determining whether ideals will in fact decline over time. Moreover, we know that individuals with high θ_i will be more prone to choose red; these individuals can also be expected to have higher β_i (Theorem 1). Thus, individuals' ideals – and thus also their contributions g_i^t – must be expected to decline faster in blue groups than in red.[6]

Assume that every i has the same exogenous *initial* ideal $\hat{g}_i^1$. Moreover, assume that after each period, individuals observe the average contribution in their own group. Although there will be random differences in β_i between members in each red, respectively blue, group, the expected level of β_i will be equal within each group type. Since individuals hold identical initial ideals, the expected development of the ideal will also be identical within each group type. Let $\hat{g}_\tau^t$ denote the expected ideal in group type $\tau = \{R, B\}$ in period t, while $\bar{g}_\tau^t$ denotes the expected average contribution in group type τ in period t.

At the beginning of period 2, when one observation has been made, the expected difference in ideals is given by

$$\hat{g}_R^2 - \hat{g}_B^2 = (1-\lambda)(\bar{g}_R^1 - \bar{g}_B^1) = \frac{1-\lambda}{\alpha}(\bar{\beta}_R - \bar{\beta}_B).$$

where $\bar{\beta}_R = E\left[\beta_i | \tau_i = R\right]$ and $\bar{\beta}_B = E\left[\beta_i | \tau_i = B\right]$ are expected average β_i in the two group types. Hence, after the second period, $\bar{g}_\tau^{t+1}$ will differ between red and blue groups both due to different ideals in period t and to different β_i's. In general,[7]

$$\begin{aligned} \hat{g}_R^{t+1} - \hat{g}_B^{t+1} &= (1-\lambda)\left(\bar{g}_R^t - \bar{g}_B^t\right) + \lambda\left(\hat{g}_R^t - \hat{g}_B^t\right) \\ &= \frac{1-\lambda}{\alpha} t(\bar{\beta}_R - \bar{\beta}_B) > 0 \end{aligned}$$

[5] If individuals are myopic but expect members of red groups to contribute more, this would reinforce the sorting of individuals with high β_i into red groups.

[6] Whether ideals will increase, be constant or decrease in either group types, depends on the level of β_i in the group.

[7] Allowing for corner solution, the equation can be written

$$\hat{g}_R^{t+1} - \hat{g}_B^{t+1} = \frac{1-\lambda}{\alpha} \sum_{s=1}^{t} \left(\bar{\beta}_R(\hat{g}_R^s) - \bar{\beta}_B(\hat{g}_B^s)\right)$$

For a proof see Appendix A.2.

where the inequality is due to Proposition 2, securing that $\bar{\beta}_R > \bar{\beta}_B$.

Thus, individuals in red groups must be expected to have, on average, both higher ideals and higher β_i values than members of blue groups. Due to the higher values of β_i, contributions will typically be higher in red groups even in the first period; for each subsequent period, the difference in contributions will increase, because of the different development of the ideal. We summarize this as a Proposition.

Proposition 3. *Given that the ideal changes according to (8), expected contributions to the local public good is higher in red groups than in blue groups in every period, and the difference is increasing over time.*

2.5 Group choice with group specific ideals

Until now, we have assumed that the group composition is kept fixed over time. Consider now the case with T periods where each individual can choose group type for each new period, such that group composition changes over time.

In the previous section, we saw that the ideal developed differently in red and blue groups. Since any given individual was then only member of one group throughout the repeated game, the ideal in the other group type would be largely irrelevant for her; as the game evolved, she might not even be aware that contributions were developing differently. Let us now assume that after each period, every individual receives information about average contributions in at least one blue and at least one red group. If the average contributions differ between group types, individual i's idea of the morally ideal contribution may not be the same for both group types: if she expects others in her group to contribute a lot, it seems more reasonable that she will demand more of herself, and vice versa.

Let us thus modify (8), the equation determining i's ideal, such that in any period $t > 1$, i has two ideals, $\hat{g}^t_{iR}$ and $\hat{g}^t_{iB}$, one for each group type, where the ideal for group type τ is updated using the information i has received about average contributions in group type τ:

$$\hat{g}^{t+1}_{i\tau} = (1-\lambda)\bar{g}^t_{i\tau} + \lambda\hat{g}^t_{i\tau} \tag{12}$$

where $\bar{g}^t_{i\tau}$ is the information received by i in period t about average contributions in group type τ. Since individuals within the same group type have different β_i's and are distributed randomly into groups, we may have $\hat{g}^t_{i\tau} \neq \hat{g}^t_{j\tau}$ for $i \neq j$ if i and j have received information from different groups. Still, an outside observer has of course no reasons initially to have different expectations for i and j. We can thus define a group type specific expected ideal: $\hat{g}^t_\tau = E(\hat{g}^t_{i\tau}) = E(\hat{g}^t_{j\tau})$.

A main difference between constant and changing group compositions over time, is that constant group composition gives a stronger strategic motive for early contributions. We

showed above, however, that due to myopia group choice did not depend on such dynamics even for constant group compositions, only on whether $D \gtrless \theta_i$. Moreover, the choice of red versus blue and the choice of g_i^t were independent. With myopic individuals, the results will be very similar when group composition is allowed to change over time. In the first period, individuals with high β_i will self-select into red groups, just as before; thus expected average contributions in the first period will be higher in red than in blue groups, the result is that $\hat{g}_R^1 > \hat{g}_B^1$. For later periods, i's group-specific ideals $\hat{g}_{i\tau}^t$ are updated according to (12); and both values of β and ideals are in expectation higher in red than blue groups. The result is that $\hat{g}_R^t > \hat{g}_B^t$, for every period. Moreover, as more periods pass, the difference must be expected to increase.

Now, let $g_{\tau i}^t$ denote the contribution of individual i in period t given that i has chosen to be in a group of type τ in this period. (As i can actually be in only one type of group at a time, either g_{Ri}^t or g_{Bi}^t will be hypothetical.) The first order conditions for a utility-maximizing contribution $g_{\tau i}^t$ is now

$$g_{\tau i}^t = \hat{g}_{\tau i}^t + \frac{1}{\alpha}\left(\beta_i - \frac{n-M}{n}\right) \tag{13}$$

Note that the last term is independent of group type. Since $\hat{g}_R^t > \hat{g}_B^t$, this Proposition follows:

Proposition 4. *For $t > 1$, individuals will in expectation contribute more as a member of a red group than as member of a blue group.*

The following Proposition also follows directly from equation (13):

Proposition 5. *Contributions are increasing in agents' ideals.*

Finally, consider the choice between red and blue. As above, those who choose red have high values of θ_i and, by correlation, of β_i. When considering ideals that depend on group type, there is an additional argument for choosing red, which is strongest for those with the highest values of β_i: Entering a red group will induce i to make larger contributions than she would in a blue group. Thus, red group membership yields more warm glow both in terms of the unconditional warm glow of global public good provision, θ_i, but also more warm glow of local public good provision, $\beta_i g_i$, which is increasing in β_i. Thus, this argument reinforces the expected self-selection of cooperative individuals into the red groups. We can summarize it in the following Proposition, for which the formal details are in Appendix A.3:

Proposition 6. *Individuals are more likely to choose red groups the higher the observed difference in contributions $\bar{g}_{iR}^t - \bar{g}_{iB}^t$ is.*

3 Experimental design

Let us now explain our experimental design in somewhat more detail. As mentioned in the introduction, the experiment consists of 3 parts, a one-shot public good game, a repeated game with fixed groups (a partner game), and a sequence of games with varying group composition (a stranger game). All subjects participate in all three parts in the same order. We denote by $t = 0$ the one-shot game, by $t = 1, \ldots, 10$ the 10 periods of the partner game, and by $t = 11, \ldots, 30$ the 20 periods of the stranger game.

Part 1 of the experiment is a standard one-shot public good game experiment with exogenous group formation. Below, we will call this part the *one-shot game*. Groups of 3 are formed randomly, and each subject is given an initial endowment of NOK 60.[8] The subject's task is to decide how much to allocate to her group and how much to keep for herself. Every contribution to the group is doubled by the experimenters, and then divided equally between the three group members. Each subject is thus paid according to the following monetary payoff function:

$$x_i^0 = 60 - g_i^0 + \frac{2}{3}\sum_{j=1}^{3} g_j^0 \tag{14}$$

where x_i^0 is subject i's monetary payoff and g_i^0 is i's contribution to her group in the one-shot game.

Before this part starts, subjects are tested in their understanding of the instructions. They are informed that two additional experiments will take place after this one, but that their choice in part 1 neither will affect their payoffs nor available choices in parts 2 and 3. No further information is provided about the contents of parts 2 and 3. Moreover, before proceeding to part 2, subjects receive no feedback about what has happened in part 1.

Part 2 of the experiment is a repeated public good game where the group composition remains fixed for all 10 periods. In what follows, this part will be called the *partner game*. The stage game is very similar to the one-shot game described above. Before the repeated public good game is played, however, each subject can choose whether she prefers to be in a red or blue group[9]. For each member of a red group, an extra NOK 50 is donated to the Norwegian Red Cross. Each member of a blue group receives, instead, the extra NOK 50 herself. To the largest extent possible, subjects are allocated to groups of their preferred type; if the number of subjects preferring a given group type is not divisible by 3, there will be one mixed group, and the type of this group is determined by the majority preference within the group.

[8] Given the exchange rate at the time of the experiment (February 2008), this was equivalent to about 11 USD.

[9] In the experiment, group types were called Z groups and X groups instead of red and blue, in order to avoid framing effects. To avoid confusion, there is a reminder of the difference on every screen where subjects make choices concerning X and Z.

The payoff function for a member of a red group in any period of the partner game (part 2) can then be written as

$$x_{iR}^t = 60 - g_i^t + \frac{2}{3}\sum_{j=1}^{3} g_j \tag{15}$$

for $t = 1, ..., 10$, and for a member of a blue group,

$$x_{iB}^t = 110 - g_i + \frac{2}{3}\sum_{j=1}^{3} g_j \tag{16}$$

for $t = 1, ..., 10$.

Actual payment to i for her choices in the partner game is determined by her average calculated payoff through all periods 1 to 10 (that is, $\frac{1}{10}\sum_{t=1}^{10} x_{iR}^t$ or $\frac{1}{10}\sum_{t=1}^{10} x_{iB}^t$, respectively). For given contribution levels, the monetary payoff to members of blue groups must obviously be higher than in red groups. However, we can have $x_{iR}^t > x_{iB}^t$ ($t \in \{1, ..., 10\}$) if the contributions of the other two group members are sufficiently higher in the red group. With our design, we will have $x_{iR}^t > x_{iB}^t$ if the average other in the red group contributes at least NOK 37,5 more than the average other in the blue group (that is, if the sum of others' contributions is at least NOK 75 higher in the red than in the blue group).

Subjects are tested in their understanding of the instructions before the partner game starts. After each period, every subject receives feedback on how many units she contributed, how many units were contributed on average in her group, and her calculated monetary payoff from that period, x_{iB}^t or x_{iR}^t. Due to the feedback given, individual choices in the partner game can be considered independent observations only in period 1.

Part 3 of the experiment is quite similar to part 2, but now new groups are formed in every period. Below, this part will thus be called the *stranger game*. Moreover, subjects now decide on their preferred group type again in each new period, and the number of periods is increased to 20.

Period payoffs x_{iR}^t and x_{iB}^t for $t = 11, ..., 30$ are calculated exactly as in the partner game, that is, according to eq. (15) and (16). Actual monetary payout to subjects is determined as the average calculated period payoff over all 20 periods (but note that now i may be a member of a red group in some periods and of a blue group in others).

Among those who have chosen the same group type, groups are formed randomly in every period. As in part 2, if the number of subjects preferring a given group type is not divisible by 3, there will be one mixed group.

Between each period, each subject receives information about how many units she contributed to the public good in that period, and how many units were contributed on average in her group. In addition, each subject is informed of the average contribution in one red

group and one blue group.[10] Due to the feedback given between periods, observations cannot be considered independent.

4 Experimental results

The experiment was programmed in z-tree (Fischbacher, 2007) and conducted at the Oeconlab at the University of Oslo in February 2008. 87 subjects recruited among students from several departments at the University of Oslo participated in the experiment.

Figures 1 and 2 summarize the data. Recall that period 0 corresponds to the one-shot game, period 1-10 to the partner game, and period 11-30 the stranger game. In figure 1 subjects' choice of group type is shown in the lower panel, while actual group types are shown in the upper panel. Remember that in sessions where the number preferring a given group type was not dividable by three, a mixed group was established, which type was decided by the majority preference of it's members.

As illustrated in the upper panel, there were 41% pure red and 52% pure blue groups, while 7% of groups were mixed in the partner game (here illustrated by the lighter shade). In the stranger game, the number of red groups was relatively stable during the 20 periods; averaged across all periods, the average percentage of subjects in red groups was 36%.

The lower panel of Figure 1 shows players' group choices. Again, the fractions remain fairly stable over the course of the game. The darker colors illustrate that about 25% of our subjects consistently chose blue, while 10% consistently chose red, during the entire 20 periods of the stranger game. The remaining 65% switched at least once between choosing red and blue during the stranger game.

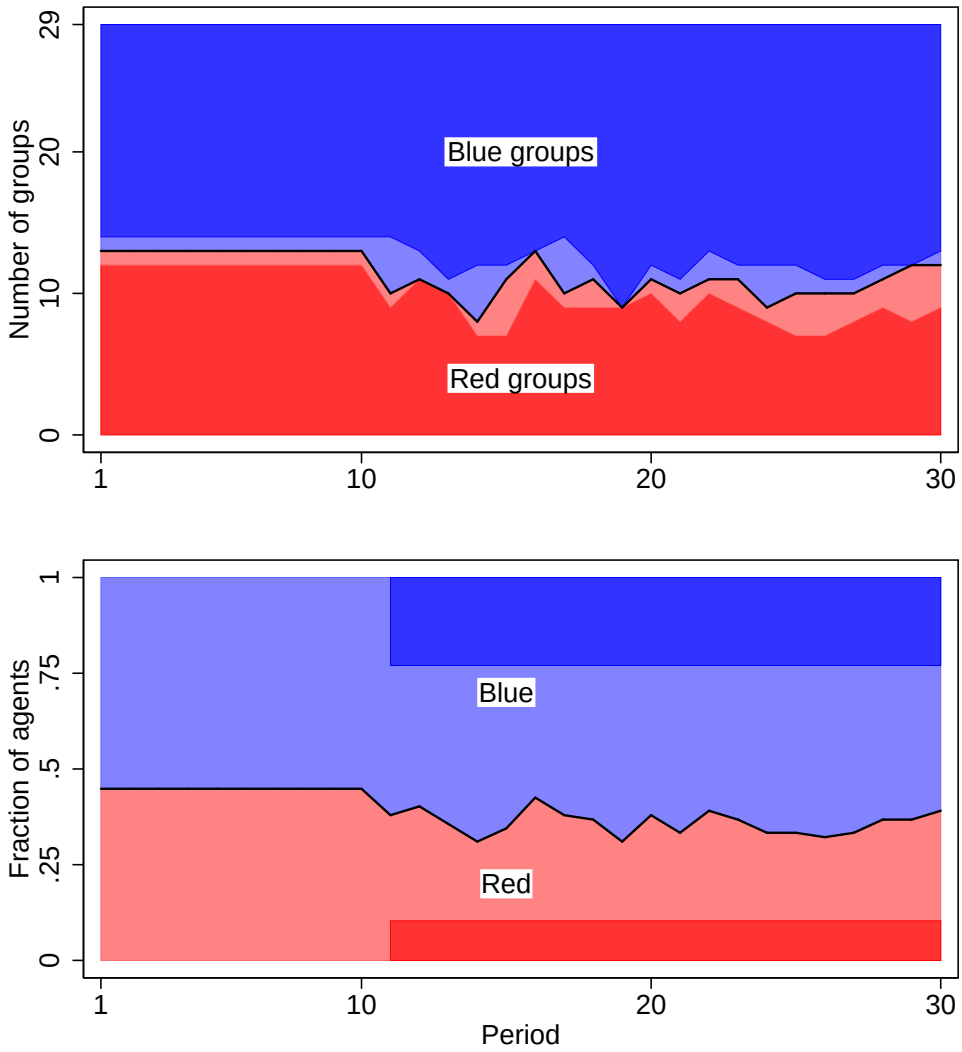
4.1 Contributions

Figure 2 illustrates average contribution levels of red and blue groups respectively. The figure reveals a rather striking pattern: First, average contribution in red groups is higher than in blue groups in the one-shot game, in every period of the partner game, and in every period of the stranger game. Moreover, contributions in the red groups do not, in contrast to the blue groups, appear to decrease over time.

Following Proposition 2, we expect those who have self-selected into red groups to contribute more than others to the local public good. We can thus formulate the following hypothesis:

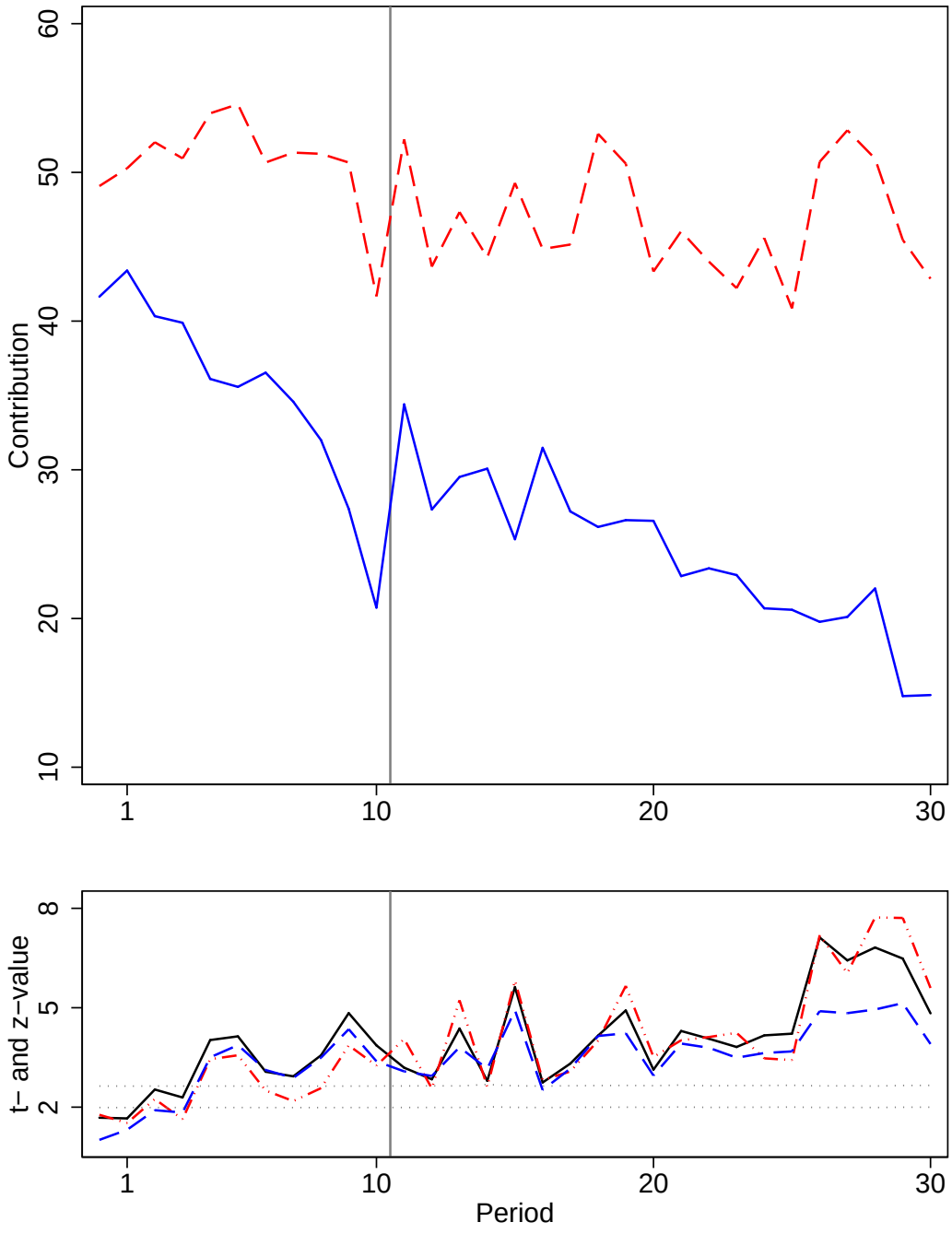
[10] To maximize the variation in the information received by different subjects, each subject is either shown the average contribution of the red group with the highest average contribution or that of the red group with the lowest average contribution, each with 50% probability. Likewise, each subject is shown the average contribution of the blue group with either the highest or the lowest average contribution, each with 50% probability. In the instructions, subjects are simply informed that they will be shown the average of one red and one blue group.

Figure 1: Group choice by period



Notes: The first panel shows the number of red and blue groups by period. Mixed groups in light colors, homogeneous groups in dark colors.
The second panel shows the fraction of agents who choose red and blue groups by period. The dark colored areas show the fraction of agents who consistently choose either red or blue groups in all the 20 periods in the stranger game.

Figure 2: Contribution in the different group types



Above: Average contribution by red and blue groups by period.
Below: Test statistics of the difference between red and blue contributions.
The solid line represents t-statistics from t-tests, the dashed-dotted line similar tests clustering at the group level, and the dashed line z-statistics from Mann-Whitney tests. The horizontal dotted lines are significance levels of 0.05 and 0.01.
Data from mixed groups are excluded.

Hypothesis 1. *In any given period, contributions to the local public good are higher in red groups than in blue groups.*

The lower panel of Figure 2 shows the test statistics from an ordinary t-test, a t-test clustering at the group level, and a Mann-Whitney test respectively. All these test whether mean contributions are equal for subjects in red and blue groups (for period 0: those who later choose red, respectively blue). The two t-tests assume draws from a normal distribution, while the Mann-Whitney test is robust to the distribution of the population, but with the cost of weaker test power. As shown in the lower panel of Figure 2, the ordinary t-test gives support for Hypothesis 1 from period 2 onward, at a 5% significance level. Since individual contributions cannot be expected to be independent between periods, however, this tells us only that contributions are significantly higher in red groups, and does not necessarily indicate that subjects in red and blue groups are drawn from different populations. The clustered t-test and the non-parametric Mann-Whitney test give support for hypothesis 1 from period 4 onwards.

Utilizing the whole panel of observations gives us further support for Hypothesis 1. This is done in Table 1, where we regress contribution on group choice. We study the partner and stranger games separately and jointly, with and without period dummies. Columns (1) and (2) show that in the partner game, contributions to the local public good is on average 15 units higher in red than blue groups. This difference is significant. Nothing much happens with the coefficient when period dummies are included. As members of one group play against each other in ten consecutive periods, we might worry about correlation between players. To solve this potential problem, all regressions are clustered at the group level.[11] In the stranger game, where players can choose group type in each period, the results are even stronger: Now, contributions in red groups are about 21 units higher than in blue groups.

Figure 3 shows the average profit in each group as well as the profit in red groups plus the contribution to the Red Cross. We see that it is never privately profitable to join a red group, so it is not rational to do so strategically.[12] However, as contributions on average are higher, the sum of private profits and contributions to the Red Cross (which corresponds to the social surplus) is higher than private profits in blue groups.

From Proposition 1, we know that contributions should decline over time in groups with sufficiently low warm glow of giving β_τ. This leads to the hypothesis

Hypothesis 2. *Contributions averaged over both group types decline over time in the partner game.*

[11] For the stranger game and the joint analysis, groups do not keep together, so we instead cluster on the group-period.

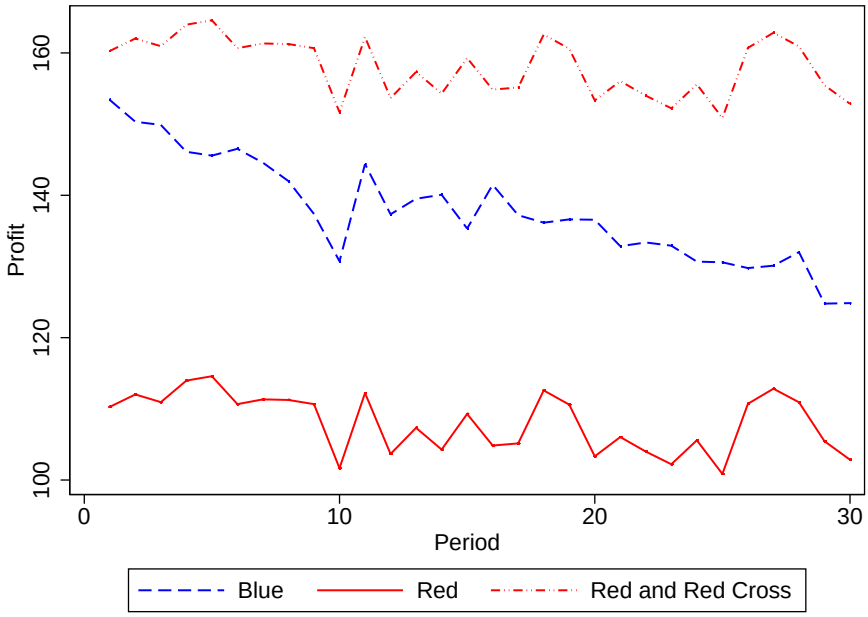
[12] Contributing 0 yields an average profit of 87 in blue groups and 62 in red groups, and the difference is positive although declining in all periods.

Table 1: Contribution by group type

	(1)	(2)	(3)	(4)	(5)	(6)
Red group	15.35***	15.35***	21.15***	21.10***	19.73***	19.09***
	(3.10)	(3.08)	(19.27)	(19.45)	(20.40)	(20.01)
Clustered at	Group	Group	Group×Period	Group×Period	Group×Period	Group×Period
Treatment	2	2	3	3	2 and 3	2 and 3
Period dummies	No	Yes	No	Yes	No	Yes
R^2	0.104	0.146	0.169	0.189	0.151	0.197
Obs	870	870	1740	1740	2610	2610

*Notes: The dependent variable is contribution. Z-values clustered at group or group-period level are reported in parenthesis. *, **, and *** denote significance at the 10%, 5%, and 1% levels.*

Figure 3: Average profit by group type



Note: Data from mixed groups are excluded.

This pattern is well-known and well-documented in the literature (Ledyard, 1995). Column (1) of Table 2 shows that this pattern is also found in our data; contributions are lower in later periods, and this difference is statistically significant.

From Figure 2, we see that this is mostly driven by the decline in the blue groups. This is in line with Proposition 3, which leads to the following hypothesis:

Hypothesis 3. *The difference in contributions between red and blue groups will increase over time.*

Column (2) of Table 2 confirms that declining contributions are, indeed, mostly found in blue groups. The interaction term between period and a dummy for being in a red group is substantially and significantly higher in the last periods, which supports Hypothesis 3.

If subjects strive towards some ideal which is positively affected by other group members' expected contributions, then the same individual will contribute (weakly) more to the local public good as a member of a red group than if she is in a blue group. From Proposition 4, we can formulate our fourth hypothesis:

Hypothesis 4. *An individual will contribute more to the local public good when member of a red group than when member of a blue group.*

Contributions in the stranger game are regressed on group choice in Table 3. As the table shows, contributions are significantly higher in red groups in all specifications. Note

Table 2: Contributions by period

	(1)	(2)		
Period	All	Blue	Red	Difference
2	-1.092	-2.875	1.103	3.978
	(-0.76)	(-1.53)	(0.53)	(1.40)
3	-2.115	-3.708	-0.154	3.554
	(-1.06)	(-1.21)	(-0.07)	(0.97)
4	-3.276	-8.500***	3.154	11.65***
	(-1.59)	(-3.02)	(1.22)	(2.83)
5	-3.299	-8.583***	3.205	11.79***
	(-1.33)	(-2.84)	(1.30)	(3.41)
6	-3.736*	-7.271**	0.615	7.886**
	(-1.83)	(-2.39)	(0.29)	(2.15)
7	-5.276**	-8.687***	-1.077	7.611*
	(-2.34)	(-2.97)	(-0.36)	(1.82)
8	-6.184***	-10.69***	-0.641	10.05**
	(-2.65)	(-3.78)	(-0.20)	(2.26)
9	-9.966***	-17.52***	-0.667	16.85***
	(-3.46)	(-4.55)	(-0.25)	(3.69)
10	-17.24***	-23.75***	-9.231**	14.52***
	(-6.22)	(-7.77)	(-2.28)	(2.84)
R2	0.146		0.156	
N	870		870	

Notes: Dependent variable is contribution. Specification (1) includes dummies for each period in the partner game. Specification (2) includes separate period dummies for members of red and blue groups, and the differences between the estimated coefficients are also reported. Z-values clustered at the group level are reported in parentheses.
, **, and * denote significance at the 10%, 5%, and 1% levels.*

Table 3: Results from the stranger game: Contributions

	(1)	(2)	(3)	(4)
Red group	21.15***	21.10***	12.46***	12.27***
	(7.52)	(7.43)	(5.23)	(5.25)
Period FE	No	Yes	No	Yes
Indivdual FE	No	No	Yes	Yes
R2	0.169	0.189	0.064	0.115
N	1740	1740	1740	1740

Notes: The dependent variable is contribution. Z-values clustered at the individual level in parentheses.
, **, and * denote significance at the 10%, 5%, and 1% levels.*

particularly the individual fixed effects model in specifications (3) and (4). An individual who changes group type during the stranger game contributes on average about 12 units more to the public good when she is a member of a red group than when she is in a blue group. This supports Hypothesis 4.

The theoretical model introduced above assumes that contributions depend on a moral ideal, which depends on the observed level of contributions in groups. This was summarized in Proposition 5, and leads us to formulate the following Hypothesis:

Hypothesis 5. *Contributions are increasing in the perceived average contribution level in one's group.*

To explore how perceptions about contributions in red and blue groups affect contributions and group choice, we gave subjects differentiated information in the stranger game. As explained above, at the end of each period subjects were informed about the contributions in *one* red and *one* blue group in addition to their own group, but the players did not get information about the same groups.

The rational way to aggregate this information is though Bayesian updating. Appendix B explains how we model this in the current context. Essentially, we assume that each subject holds a prior distribution of the mean contribution among red and blue group members. After each period, subjects use observations from their own group, as well as the provided information, to update their beliefs according to Bayes' rule. The calculated beliefs are then used as explanatory variables in the estimations shown in Tables 4 and 5. Here, $\mu_{i\tau}^t$ denotes player i's best estimate of contributions in group type τ in period t, and $\Delta\mu_i^t = \mu_{iR}^t - \mu_{iB}^t$ is her expected difference in contributions between red and blue groups.

Table 4 shows the effect of beliefs on own contribution. Columns (1) and (2) show results from the whole sample in the stranger game, with and without individual fixed effects. In both specifications, increased belief about others' contributions increase own contributions,

Table 4: Results from stranger game: Effect of beliefs

	(1)	(2)	(3)	(4)
Red group	20.19***	12.43***		
	(7.20)	(5.03)		
μ_{blue}	0.0216	0.216**	0.294**	0.00479
	(0.10)	(2.45)	(2.17)	(0.07)
μ_{red}	0.690***	0.123	0.0780	0.321**
	(3.32)	(1.08)	(0.55)	(2.20)
Period FE	Yes	Yes	Yes	Yes
Indivdual FE	No	Yes	Yes	Yes
Group type	Both	Both	Blue	Red
R2	0.216	0.109	0.067	0.051
N	1653	1653	1060	593

Notes: The dependent variable is contribution. The μ's are from the Bayesian updating, i.e. the best guess of the mean contribution of blue and red players given the subject's available information. Z-values clustered at the individual level in parentheses.
, **, and * denote significance at the 10%, 5%, and 1% levels.*

although not always significantly. In columns (3) and (4), regressions are run separately for red and blue groups. Now, belief about the group one is in has a positive and significant effect, whereas the belief of the other group has no significant effect. All in all, these findings tend to support Hypothesis 5.

4.2 Group choice

In the experiment, players can choose what type of group to join: before the partner game, and before each period of the stranger game. As a parallel to Proposition 2 we should expect that players that had high contributions in the one-shot game should have a tendency to choose red groups later on. As this is purely a question of correlation and not a causal relationship, we can also study this by testing whether contributions are higher in the one-shot game for players who chose red at the beginning of the partner game than those who chose blue. From the upper panel of Figure 2, we see that contributions are actually higher in the former group, but the lower panel reveals that this difference is not statistically significant.

More interesting is the sequence of decisions in the stranger game. These are analyzed in Table 5. Column (1) is a simple estimation to indicate persistence in group choice: Having chosen a red group in one period substantially and significantly increases the probability of choosing a red group in the next period. As a counterpart to Proposition 6 in the theoretical section, we have the following hypothesis regarding group choice:

Table 5: Results from stranger game: Group choice

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Group choice	2.706***	2.707***	2.744***	2.739***	2.677***	2.005***	0.441**
	(21.24)	(21.20)	(20.95)	(20.90)	(20.97)	(14.60)	(2.12)
$\Delta\mu$		0.0211***		0.00791		0.0543***	0.0106
		(2.62)		(0.90)		(3.94)	(0.65)
Δ Shown			0.0134***	0.0120***			
			(4.09)	(3.38)			
μ_blue					0.00205		
					(0.18)		
μ_red					0.0452***		
					(3.82)		
$\Delta\mu\times$ Red in partner game						-0.0417**	
						(-2.38)	
Red group in partner game						2.505***	
						(7.50)	
$\Delta\mu\times$ Red in stranger game							0.0334
							(1.17)
Red group in stranger game							3.944***
							(7.60)
ρ^2	0.259	0.263	0.265	0.265	0.267	0.342	0.396
N	1653	1653	1608	1608	1653	1653	1653

Notes: Estimates are logit coefficients with choice of red type next period as the positive outcome.
The μ's are from the Bayesian updating, i.e. the best guess of the mean contribution of blue and red players.
The variable Δ Shown is the difference in the contributions actually shown.
Red group in stranger game is the fraction of time so far in the stranger game spent in red groups.
All specifications include period dummies.
ρ^2 is McFadden's goodness of fit measure. Robust Z-values are reported in parentheses
, **, and * denotes significance at the 10%, 5%, and 1% levels.*

Hypothesis 6. *The probability of choosing a red group is increasing in the perceived difference between mean contributions in red and blue groups.*

From the Bayesian updating, we have a proxy for the difference in beliefs, $\Delta\mu_i^t = \mu_{iR}^t - \mu_{iB}^t$. Column (2) of Table 5 is a regression where we only include this estimated belief about the difference of the mean contributions between red and blue groups. We can see that higher expected contributions in red as compared to blue groups tend to increase the likelihood of choosing a red group in the next period. Given that the potential range of $\Delta\mu$ is -60 to 60 and the observed range -31 to 45, the quantitative effect is substantial. In Column (3) we instead include the difference between the two pieces of information provided in the current period, which has a similar effect as the μs. In Column (4) we set the two up against each other, and here it turns out that the current information retains significance whereas $\Delta\mu$ becomes insignificant. Hence it seems as players put more emphasis on recently received information. To see whether information about both groups are equally important, we replace $\Delta\mu$ with the separate μs in Column (5). In sum, there seems to be support for Hypothesis 6, but agents are more short sighted than we would expect from theory.

Finally, it is interesting to explore whether subjects in red groups react differently to updated information about others behavior than those in blue groups. Column (6) of Table 5 interacts the belief variable with an indicator variable for having chosen a red group in the partner game. We see that subjects from red groups in the partner game, do, in fact, react *less* strongly to information. In Column (7), we undertake a similar estimation, this time with information interacted with the fraction of time spent in red groups so far in the stranger game. It seems that this variable has little impact on information processing.

5 Conclusion

In public good games, many subjects contribute relatively large amounts in the first period, but decrease their contributions substantially as the game is repeated. In this paper we have proposed a formal model of this, based on the assumption that both unconditional and conditional altruism play a role. Using the same formal framework, we have shown that if there is a positive correlation between the strength of individuals' unconditional altruism towards different public goods, subjects who are highly cooperative will self-select into groups committed to charity donations, and these groups may be able to sustain cooperation.

Our predictions were tested by means of a public good game with endogenous group formation. Before making their contribution choice, subjects could choose which type of group they preferred to be a member of. The difference between group types was that a fixed extra amount of money was donated either to the Red Cross (red groups) or to individual subjects (blue groups). Groups were then formed randomly, but to the largest

extent possible respecting subjects' group type preferences.

Throughout the periods, the share of red groups was always between 30 and 45 percent, with no clear trend in either direction. Average contributions were significantly higher in red groups. The difference was not sufficiently large, however, to leave red group membership equally profitable, in monetary terms, as blue group membership. Thus, individuals hardly chose red group membership in the hope of earning more money due to others' high contributions.

High contributors in an independent one-shot public good game also chose red groups more often than low contributors, although not significantly so. However, we found that while contributions in blue groups decreased substantially over time, contributions stayed high in the red groups. Finally, we found that subjects who shifted between being in red and blue groups contributed significantly more when in a red group compared to when in a blue group.

In line with previous experimental evidence, we thus have shown both theoretically and empirically that it is possible to sustain cooperation provided that the group consists of the right people: "the good guys". In addition, we have shown that pre-commitment to charitable giving is one possible way for these good guys to find each other.

Willingness to contribute to one public good (here, the Red Cross) can be taken as a signal of high willingness to contribute to other public goods as well (here, the group's material benefit). Committing oneself to a group donating to charity consequently increases the chance of being matched with others who contribute a lot to the group. Thus matched, conditional cooperators can sustain high average contribution levels. While a pre-commitment to charity donations tend to attract high contributors, the increasing difference in average contributions over time in the two group types is driven by the dynamic interaction between conditional cooperators. The question is therefore not only whether you are a good or a bad guy; but, perhaps more importantly, whether you are playing with the good guys.

References

Ahn, T., R. M. Isaac, and T. C. Salmon (2008): "Endogenous group formation," *Journal of Public Economic Theory*, 10, 171–194.

Altmann, S., T. Dohmen, and M. Wibral (2008): "Do the Reciprocal Trust Less?," *Economics Letters*, 99, 454–57.

Andreoni, J. (1988a): "Privately Provided Public Goods in a Large Economy: The Limits of Altruism.," *Journal of Public Economics*, 35(1), 57–73.

Andreoni, J. (1988b): "Why free ride?," *Journal of Public Economics*, 37, 291–304.

Andreoni, J. (1990): "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *Economic Journal*, 100(401), 464–477.

Blanco, M., D. Engelmann, and H. Normann (2007): "A Within-Subject Analysis of Other-Regarding Preferences," Mimeo, Royal Holloway, University of London.

Brekke, K. A., and K. Nyborg (2008): "Attracting Responsible Employees: Green Production as Labor Market Screening," *Resource and Energy Economics*, 30, 509–526.

Brekke, K. A., K. Nyborg, and S. Kverndokk (2003): "An economic model of moral motivation," *Journal of Public Economics*, 87, 337–3354.

Croson, R. T. A. (1996): "Partners and strangers revisited," *Economics Letters*, 53(1), 25–32.

Ehrhart, K.-M., and C. Keser (1999): "Mobility and cooperation: On the run," Working Paper 99s-24, CIRANO, Montreal.

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

——— (2002): "Altruistic punishment in humans," *Nature*, 415(6868), 137–140.

Festinger, L. (1957): *A Theory of Cognitive Dissonance*. Stanford University Press, Stanford, CA.

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

Fischbacher, U., and S. Gächter (2006): "Heterogeneous social preferences and the dynamics of free riding in public goods," IZA Discussion Papers 2011.

——— (forthcoming): "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods," *American Economic Review*.

Fischbacher, U., S. Gächter, and E. Fehr (2001): "Are people conditionally cooperative? Evidence from a public goods experiment," *Economics Letters*, 71, 397–404.

Gächter, S., and C. Thöni (2005): "Social learning and voluntary cooperation among like-minded people," *Journal of the European Economic Association*, 3(2-3), 303–314.

Gunnthorsdottir, A., D. Houser, and K. McCabe (2007): "Disposition, history and contributions in public goods experiments," *Journal of Economic Behavior and Organization*, 62(2), 304–315.

Gürerk, Ö., B., and Rockenbach (2006): "The competetive Advantage of Sanctioning Institutions," *Science*, 312, 108–111.

Hauge, K. E. (2007): "Normative reflection in a public good game experiment," Mimeo, University of Oslo.

Holmström, B. (1982): "Moral Hazard in Teams," *Bell Journal of Economics*, 13, 324–340.

Johnson, N. L., and S. Kotz (1972): *Distributions in Statistics: Continuous Multivariate Distributions*. John Wiley & Sons, New York.

Johnson, T., C. T. Dawes, J. H. Fowler, R. McElreath, and O. Smirnov (2009): "The role of egalitarian motives in altruistic punishment," *Economics Letters*, 102, 192–94.

Keser, C., and F. van Winden (2000): "Conditional Cooperation and Voluntary Contributions to Public Goods," *Scandinavian Journal of Economics*, 102(1), 23–39.

Konow, J. (2009): "Mixed Feelings: Theories and Evidence of Altruism," Mimeo, Loyola Marymount University.

Ledyard, J. O. (1995): "Public goods: A Survey of Experimental Research," in *The Handbook of Experimental Economics*, ed. by J. H. Kagel, and A. E. Roth, pp. 111–194. Princeton University Press, Princeton, New Jersey.

Neugebauer, T., J. Perote, U. Schmidt, and M. Loos (2009): "Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments," *Journal of Economic Psychology*, 30, 52–60.

Ones, 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(4), 495–521.

Page, T., L. Putterman, and B. Unel (2005): "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency," *The Economic Journal*, 115(October), 1032–1053.

Sugden, R. (1984): "Reciprocity: The supply of Public Goods Through Voluntary Contributions," *The Economic Journal*, 94(dec), 1281–1302.

A Proofs

A.1 Proof of Theorem 1

We have $(b_i, y_i)' \sim N\left[\left(\mu_b, \mu_y\right)', \Sigma\right]$. Then it follows from standard results on conditional normal distributions (see e.g. Johnson and Kotz, 1972, Ch 35.3) that

$$b_i \left| y_i = Y \right. \sim N\left[\left(\mu_b + \frac{\sigma_{12}}{\sigma_2^2}\left(Y - \mu_y\right)\right), \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}\right]$$

where

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

Hence $b_i \left| y_i > Y \right.$ has a distribution function

$$f_Y^+(b) = \int_{x=Y}^{+\infty} \phi\left(\frac{b - \left(\mu_b + \frac{\sigma_{12}}{\sigma_2^2}\left(x - \mu_y\right)\right)}{\sqrt{\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}}}\right) \frac{\phi\left(\frac{x-\mu_y}{\sigma_2^2}\right)}{1 - \Phi\left(\frac{Y-\mu_y}{\sigma_2^2}\right)}\, dx$$

whereas $b_i \left| y_i < Y \right.$ has a distribution function

$$f_Y^-(b) = \int_{x=-\infty}^{Y} \phi\left(\frac{b - \left(\mu_b + \frac{\sigma_{12}}{\sigma_2^2}\left(x - \mu_y\right)\right)}{\sqrt{\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}}}\right) \frac{\phi\left(\frac{x-\mu_y}{\sigma_2^2}\right)}{\Phi\left(\frac{Y-\mu_y}{\sigma_2^2}\right)}\, dx$$

where ϕ and Φ denote the pdf and cdf of the standard normal distribution. Hence the former stochastically dominates the latter if $\int_{-\infty}^{z} f_Y^-(b) - f_{-\infty}^-(b)\ db > 0$ for all z, which holds when

$$\int_{x^+=Y}^{+\infty} \int_{x^-=-\infty}^{Y} \int_{b=-\infty}^{z} \left[\phi\left(\frac{b - \left(\mu_b + \frac{\sigma_{12}}{\sigma_2^2}\left(x^- - \mu_y\right)\right)}{\sqrt{\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}}}\right) - \phi\left(\frac{b - \left(\mu_b + \frac{\sigma_{12}}{\sigma_2^2}\left(x^+ - \mu_y\right)\right)}{\sqrt{\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}}}\right)\right] db \tag{17}$$

$$\times \frac{\phi\left(\frac{x-\mu_y}{\sigma_2^2}\right)}{1 - \Phi\left(\frac{Y-\mu_y}{\sigma_2^2}\right)} \frac{\phi\left(\frac{x-\mu_y}{\sigma_2^2}\right)}{\Phi\left(\frac{Y-\mu_y}{\sigma_2^2}\right)}\, dx^- dx^+ > 0.$$

As we always have $x^- < x^+$,

$$\int_{b=-\infty}^{z} \phi\left(\frac{b - \left(\mu_b + \frac{\sigma_{12}}{\sigma_2^2}\left(x^- - \mu_y\right)\right)}{\sqrt{\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}}}\right) - \phi\left(\frac{b - \left(\mu_b + \frac{\sigma_{12}}{\sigma_2^2}\left(x^+ - \mu_y\right)\right)}{\sqrt{\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}}}\right)\ db > 0.$$

It follows that (17) always holds, and hence that $b_i|y_i > Y$, will stochastically dominate $b_i|y_i < Y$. It trivially follows that we can replace y_i and Y with $\theta_i = g(y_i)$ and $\hat{\theta} = g(Y)$ for any monotonically increasing function g, so $b_i|$red stochastically dominates $b_i|$blue. Finally stochastic dominance is preserved by monotonically increasing relations, so $\beta_i|$red stochastically dominates $\beta_i|$blue for $\beta_i = h(b_i)$ for any monotonically increasing function h.

A.2 Proof of footnote 7

Suppose that the claim

$$\hat{g}_R^{t+1} - \hat{g}_B^{t+1} = \frac{1-\lambda}{\alpha} \sum_{s=1}^{t} (\bar{\beta}_R(\hat{g}_R^s) - \bar{\beta}_B(\hat{g}_B^s))$$

holds for $t+1 \leq k$. We want to show that it holds for $t+1 = k+1$ as well. Note first that the average contribution is given by

$$\bar{g}_\tau^t = \hat{g}_\tau^t + \frac{\bar{\beta}_R(\hat{g}_R^t)}{\alpha} - \frac{n-M}{n\alpha} \text{ for } \tau \in \{R, B\}$$

so

$$\bar{g}_R^t - \bar{g}_B^t = \left(\hat{g}_R^t - \hat{g}_B^t\right) + \frac{\bar{\beta}_R(\hat{g}_R^t) - \bar{\beta}_B(\hat{g}_B^t)}{\alpha}$$

Combining the two equations above with equation (8), we find that

$$\begin{aligned}
\hat{g}_R^{k+1} - \hat{g}_B^{k+1} &= (1-\lambda)\left(\bar{g}_R^k - \bar{g}_B^k\right) + \lambda\left(\hat{g}_R^k - \hat{g}_B^k\right) \\
&= (1-\lambda)\left(\left(\hat{g}_R^k - \hat{g}_B^k\right) + \frac{\bar{\beta}_R(\hat{g}_R^k) - \bar{\beta}_B(\hat{g}_B^k)}{\alpha}\right) + \lambda\left(\hat{g}_R^k - \hat{g}_B^k\right) \\
&= \frac{1-\lambda}{\alpha}(\bar{\beta}_R(\hat{g}_R^k) - \bar{\beta}_B(\hat{g}_B^k)) + \left(\hat{g}_R^k - \hat{g}_B^k\right) \\
&= \frac{1-\lambda}{\alpha} \sum_{s=1}^{k} (\bar{\beta}_R(\hat{g}_R^s) - \bar{\beta}_B(\hat{g}_B^s))
\end{aligned}$$

Since we know from the text that the claim is true for $k = 2$, the general claim follows by induction.

A.3 Details of group choice when ideals depend on group type

Utility will be

$$u_i^t = \begin{cases} Y + \theta_i + \frac{M}{n}(n-1)\bar{g}_R^t + (\beta_i - \frac{n-M}{n})g_i - \frac{\alpha}{2}(g_i^{t+1} - \hat{g}_{iR}^{t+1})^2 & \text{if} \quad \tau_i = R \\ Y + D + \frac{M}{n}(n-1)\bar{g}_B^t + (\beta_i - \frac{n-M}{n})g_i - \frac{\alpha}{2}(g_i^{t+1} - \hat{g}_{iB}^{t+1})^2 & \text{if} \quad \tau_i = B \end{cases}$$

Inserting the optimal contribution

$$g_i^t = \hat{g}_{i\tau}^t + \frac{1}{\alpha}\left(\beta_i - \frac{n-M}{n}\right)$$

we see that i will chose red if

$$\begin{aligned}\theta_i &\geq D + \frac{M}{n}(n-1)\left(\bar{g}_{iB}^t - \bar{g}_{iR}^t\right) + \left(\beta_i - \frac{n-M}{n}\right)\left(g_{iB}^t - g_{iR}^t\right) \\ &= D - \frac{M}{n}(n-1)\left(\bar{g}_{iR}^t - \bar{g}_{iB}^t\right) - \left(\beta_i - \frac{n-M}{n}\right)\left(\hat{g}_{iR}^t - \hat{g}_{iB}^t\right).\end{aligned}$$

Note that individual i is more likely to choose red the higher the difference in contributions $(\bar{g}_{iR}^t - \bar{g}_{iB}^t)$ is, as claimed in Proposition 6.

The condition can also be rewritten as

$$\theta_i + \beta_i\left(\hat{g}_{iR}^t - \hat{g}_{iB}^t\right) \geq D - \frac{M}{n}(n-1)\left(\bar{g}_R^t - \bar{g}_B^t\right) + \frac{n-M}{n}\left(\hat{g}_{iR}^t - \hat{g}_{iB}^t\right)$$

In expectations $\hat{g}_R^t > \hat{g}_B^t$, and thus we see that both a high θ_i as well as a high value of β_i increase the likelihood that individual i will choose the red group. We can thus expect a high value of β_i in red groups both because of the correlation with θ_i but also due to the direct impact of β_i.

The intuitive reason why β_i enters directly here is that those with a high β_i have an additional motive to enter a group where they will be expected to make high contribution, and that is the warm glow they derive from high contributions. This only matters when expected contributions are different between different group types.

B Details of the Bayesian updating

In the stranger game, agents are allowed to choose their group at the beginning of each period. Furthermore, at the end of each period, they are informed about the average contribution and profit of one blue and one red group. Assume that the contribution of a randomly chosen individual j is distributed

$$g_j|\tau \sim N(\theta_\tau, \sigma^2) \quad \tau \in \{R, B\}$$

where the variance σ^2 is assumed known but the means θ_R and θ_B are only imperfectly known. Specifically, it is assumed that at time t, agent i has prior distributions

$$\theta_\tau\,|\mathcal{F}_{i,t} \sim N(\mu_{i,\tau,t}, \omega_{i,\tau,t}^2) \quad \tau \in \{R, B\}$$

where $\mathcal{F}_{i,t}$ is i's information set at time t.

Each individual observes his own group, one blue group Γ_B and one red group Γ_R. He uses Bayesian updating to construct posterior distributions. Let $\bar{e}_i$ denote the average contribution in i's own group excluding his own contributions, and $\bar{e}_\tau$ denote the averages in Γ_τ. Finally, there are N members in both i's group in the groups Γ_τ. Let $I_{i,\tau}$ take the value 1 if i is in a group of type τ and 0 otherwise. Then the updated means are given by

$$\mu_{i,\tau,t+1} = \frac{\frac{1}{\omega^2_{i,\tau,t}}\mu_{i,\tau,t} + \frac{N}{\sigma^2}\bar{e}_\tau + I_{i,\tau}\frac{N-1}{\sigma^2}\bar{e}_i}{\frac{1}{\omega^2_{i,\tau,t}} + \frac{N}{\sigma^2} + I_{i,\tau}\frac{N-1}{\sigma^2}} \tag{18}$$

and the variances

$$\omega^2_{i,\tau,t+1} = \left(\frac{1}{\omega^2_{i,\tau,t}} + \frac{N}{\sigma^2} + I_{i,\tau}\frac{N-1}{\sigma^2}\right)^{-1} \tag{19}$$

In the empirical implementation, σ^2 is taken as known and estimated as the variance of all contributions in the stranger game. The prior $\mu_{i,\tau,0}$ is given by the average contribution of the group type, but with a high variance $\omega^2_{i,\tau,0} = 100\sigma^2$. As a consequence, for $t > 0$, $\mu_{i,\tau,t}$ is almost independent of the prior.

C Instructions

Welcome to this experiment in economics. The results from this experiment will be used in a research project. Therefore, it is important that you follow certain rules. It is important that you do not talk or in other ways communicate with any of the other participants during the experiment. Please turn off mobile phones, and use only pre-opened software on the computer. In the experiment, there will be full anonymity, which means that no other participants in this room will know which decisions you in particular make during the experiment. In addition, it is not possible to track the decisions made during the experiments back to individuals. You will be notified when the experiment starts, and when you can start entering your answers on the computer in front of you. If you have any questions during the experiment, please raise your hand, and an experimenter will come to you and answer your question in private.

You will receive money in compensation for participating in this experiment. How much money you receive will depend on the decisions you make during the experiment. After the experiment is over, you will be informed about your total payment. The person who organizes the practical payments will not now the details of the experiment and can therefore not know which decisions you have made.

Instructions Part 1

This experiment consists of three parts. Your choices in part 1 will not influence what happens, or what payment you can receive from part 2 and 3.

In part 1 of the experiment, you will be part of a group consisting of 3 people (yourself and 2 others). All three members of the group will receive an endowment of 60 NOK each. Your task is to decide how you want to allocate the money. You shall choose how many NOK you want to keep, and how much you want to contribute to an account which belongs to your group.

Your compensation for participating in the experiment, depends on how much of the endowment you choose to keep, how much you contribute to the group account, and how much the others in your group contribute to the group account. When all members of your group have decided how they wish to contribute to the group account, the total amount contributed to your group's account will be doubled and then divided equally between the three of you. For each NOK you keep, you (and only you) will earn 1 NOK. For each NOK you contribute to the group account, you and all the others in your group will earn 2/3 NOK. The same applies for the others in your group.

Examples: If for instance you contribute all your 60 NOK to the group account and the others in your group keep all for themselves, you will be paid 40 NOK (2/3 x the 60 NOK you contributed to the group account =40), while the two others will be paid 100 NOK each (the 60 NOK they kept + 2/3 x the 60 NOK you contributed to the group account). If all three group members contribute the entire endowment to the group account, each group member will be paid 120 NOK each (2/3 x 3 x 60 =120 NOK). If all three group members contribute 30 NOK, each group member will be paid 90 kroner each (the 30 NOK kept, plus 2/3 x 3 x 30 NOK = 60 NOK to each from the group account). If all three contribute zero to the group account, all three are paid the initial endowment of 60 kroner.

Notice that what happens in your group does not influence participants in other groups. Likewise, the decisions of participants in other groups than your own does not influence you.

It is important for the results of the experiment that there are no misunderstandings of the instructions. To ensure that the instructions are clear, we ask you to fill in the question sheet on the desk in front of you. This is not a test of your knowledge, but insurance for us that we have given you clear instructions. You will now get a couple of minutes to read through the instructions and answer the questions on the sheet. Raise your hand when you are finished, or if you have any questions.

Instructions Part 2

The experiment will now continue. Your decisions in part 2 will not influence what happens or the payment you can receive in part 3.

Part 2 of the experiment is quite similar to part 1. In difference from part 1, there are now first choose the type of group you prefer being a member of: X or Z.

When all participants have chosen their preferred group type, the computer will randomly create groups of 3 according to preferred group type. If the number of participants preferring one type is not divisible by three, there will be one mixed group. The type of the mixed group will be decided by the majority wish of the mixed group. All participants in all groups will be informed about what kind of group they are members of, and whether it is a homogenous or a mixed group.

Part 2 consists of 10 periods. Your actual payment from part 2 will consist of your average payoff across these 10 periods.

The group composition will be the same in all periods, and your group will NOT consist of the same individuals as in part 1 of the experiment. In each period each participant will receive an endowment of 60 NOK. Just as in part 1, your task is to decide how many NOK to keep, and how many NOK to contribute to the group account. After each period, you will receive feedback concerning how many units you yourself contributed to the group account, how many units the other two in your group on average contributed to the group account, and your calculated payoff from that period. Just like before, the total amount in the group account will be doubled, and then divided equally between the three group members. Your actual payoff from part 2 will be the average of your calculated payoffs from the 10 periods of part 2.

What is the difference between the group types? Group type X: Those organizing the experiment have inserted an extra amount of money into the group account in advance, such that each participant in every X-group will receive an extra payoff of 50 NOK per period Group type Z: Here there is no extra payoff to the participants. Instead, those that organize the experiment in each period for each participant in each Z-group will reserve 50 NOK to the Norwegian Red Cross. Just as your payoff is the average of your calculated payoff over the 10 periods, the actual payment to the Red Cross will be equal to the average of the reserved for the Red Cross over the 10 periods.

Again we ask you to fill in a sheet of questions. The sheet will be handed out, and you will get some minutes to read through the instructions on your own, and answer the questions. Raise your hand when you are finished, or if you have any questions.

Instructions Part 3

In this last part of the experiment there are 20 periods. In this part you will be able to choose group type between every period. In each period you will be part of a new group. Otherwise the rules are as they were in part 2.

After every period you will receive feedback about your calculated payoff from the previ-

ous period, and how many units were contributed to the group account on average in your group. You will also be told how many units on average was contributed to the group account in one X- and one Z-group, and the average calculated payoff for the members of these two groups. The rules for payoff are as before: Contributions from group members are doubled and then divided equally between the group members. In X-groups each member will in addition receive a calculated payoff of 50 kroner each in each period. In Z-groups 50 kroner for each member in each period is reserved for the Red Cross. From part 3 your actual payoff is equal to the average of your calculated payoff in the 20 periods. The Norwegian Red Cross will receive an amount equal to the average of what is reserved over the 20 periods.