8-29-2016

Do Economic Inequalities Affect Long-Run Cooperation?

Gabriele Camera
Chapman University, camera@chapman.edu

Cary Deck
Chapman University, deck@chapman.edu

David Porter
Chapman University, dporter@chapman.edu

Do Economic Inequalities Affect Long-Run Cooperation?

Comments
Working Paper 16-18

Do Economic Inequalities Affect Long-Run Cooperation?*

Gabriele Camera
Chapman University
University of Basel

Cary Deck
University of Arkansas
Chapman University

David Porter
Chapman University

August 29, 2016

Abstract

Does inequality affect a group's cohesion and ability to prosper? Participants in laboratory economies played an indefinite sequence of helping games in random, anonymous pairs. A coin flip determined donor and recipient roles in each pair. This random shock ensured equality of opportunity but not of results, because earnings depended on realized shocks. We manipulated the ability to condition choices on this uncontrollable inequality source. In all treatments, uncertain ending supports multiple Pareto-ranked equilibria, including full cooperation. Theoretically, inequalities do not alter the incentives' structure. Empirically, inequality disclosures altered conduct, weakened norms of mutual support and reduced efficiency.

Keywords: experiments, indefinitely repeated games, social norms, social dilemmas.
JEL codes: C70, C90, D03, E02

1 Introduction

Inequality looms large in the mind of Americans (NYT, 2015) and of people in many other countries (Pew Research Center, 2014). Social scientists' main concern is that inequality may undermine the long-run prosperity of a nation

* We thank Nat Wilcox for helpful conversations, Kladji Bregu for help running the experiments and seminar participants at Chapman University and the Cleveland Fed. G. Camera acknowledges partial research support through the NSF grant CCF-1101627.

Correspondence address: Gabriele Camera, Economic Science Institute, Chapman University, One University Dr., Orange, CA 92866; e-mail: camera@chapman.edu.

(Stiglitz, 2012). One set of problems is due to the economic inefficiency it might induce by creating distortions in capital and labor markets (e.g., Aghion and Williamson, 1998; Piketty, 2014). But inequality has also been argued to erode cohesion, trust, and cooperation (Putnam, 2000), institutions that are pillars of prosperity because they influence the structure of incentives in the economy (Kimbrough et al., 2008; North, 1991). Much less is known about this second aspect. This study contributes to filling this gap by means of an experiment.

We ask: does inequality affect cooperation and a group's ability to prosper? Field data offers ambiguous evidence on this point because many institutional and environmental factors co-vary with inequality. Prosperity may reflect changes in the capital market's structure, not wealth inequality; reduction in social cohesion may stem from migration, not income gaps; and so on. Inequality itself may stem from a mix of factors (choice, luck, power, ability). In the laboratory, these kinds of confounding factors can be controlled.

In our experiment, a group of four subjects plays an indefinite sequence of helping games as strangers, in ever-changing pairs. This helping game offers the simplest setup to think about repeated cooperation because the game is an individual decision problem. In each pair one party (the donor) has the option to cooperate, suffering a small cost to provide a larger benefit to the other party (the recipient). A virtual coin flip determines who gets which role. This random role assignment amounts to an uncontrollable shock to earning opportunities, which guarantees equal future opportunity but not equal results, as the realized sequences of shocks are inherently heterogeneous. This set-up gives rise to a social dilemma. Per-capita income increases with cooperation, so full cooperation maximizes expected payoffs. However, each donor has a short-run temptation to free ride.

In this environment, simple cooperative strategies such as "tit-for-tat" cannot be used because players do not interact as partners. However, folk theorem-type results have shown that strangers can exploit the uncertain ending, and the actions known to have been taken in their group, to remove opportunistic temptations (Kandori, 1992). Many outcomes are consistent with Nash equilibrium—including full cooperation, which is an equilibrium in which the exogenous shocks induce income inequality. Theoretically, inequality should have no effect on the structure of incentives for payoff-maximizing players because past shocks do not influence the expected return from cooperation. However, previous experiments with a similar design have shown that subjects seldom coordinate on the efficient equilibrium (Camera and Casari, 2014; Camera et al., 2013).

Our first contribution is to provide evidence that inequality that theoretically does not alter the structure of incentives, in practice influences patterns of cooperation and aggregate outcomes. The random variation in realized earning opportunities affected individual behavior and aggregate efficiency. Donors conditioned their choices on their own history of roles, rather than on the actions known to have been taken in their group—as theory suggests should happen to support the efficient outcome. The most advantaged subjects—who experienced frequent opportunities to benefit from cooperation—cooperated more than those at the other end of the spectrum. Inequality thus presents an obstacle to coordinating on efficient play when a social dilemma is indefinitely repeated.

As a second contribution, we demonstrate that cooperation is elastic to information about inequality. In a treatment, donors saw the counterparts' frequency of shocks before making a choice. Behaviorally, this makes inequality salient. Theoretically, this neither removes any of the equilibria that can be attained in the baseline set-up, nor alters the expected return from cooperation or the theoretical structure of incentives. However, we find that group cohesiveness suffered and efficiency declined compared to the baseline set-up where histories of shocks remained hidden.

Three features set our design apart from other experiments about cooperation and inequality. First, not only is full cooperation a Nash equilibrium, but it is only one of many possible equilibria. By contrast, the inefficient outcome is typically the only equilibrium in previous experimental designs (e.g., Andreoni and Varian, 1999). Second, subjects interact as strangers, not partners as in previous designs (e.g., Nishi et al., 2015). This means that cooperation hinges on the group being able to develop norms of community punishment, since individual identities and past conduct always remain hidden—which prevents reciprocity or reputation-building. Finally, cooperation has the same expected return for all group members, unlike in previous experiments (e.g., Gangadharan et al., 2015). Section 2 discusses in more detail the related experimental literature. Section 3 describes the design. Section 4 presents the theory. Section 5 reports the main results and Section 6 offers some final considerations.

2 Related studies

Our work is mainly related to experimental studies of cooperation in repeated social dilemmas and, in particular, to indefinitely repeated dilemmas—which support a richer set of equilibria compared to games that are one-shot or with a commonly known number of periods (Palfrey, 1994).

In the typical experimental design of indefinitely repeated dilemmas, the matching protocol involves fixed pairs of subjects, who take an action in every period, in a symmetric game (e.g., Blonski et al., 2011; Dal Bó and Fréchette, 2011; Duffy and Ochs, 2009; Sherstyuk et al., 2013). That design not only allows for reciprocity mechanisms to support cooperation, but it also ensures that earnings are equal under full cooperation. Recent experiments have considered a design that rules out reciprocity, and where full cooperation does not guarantee equal earnings (Camera and Casari, 2014; Camera et al., 2013). In their strangers design, pairs are randomly re-matched each period and subjects do not take an action in every period; they either have the opportunity to give a benefit to someone else, or to receive a benefit from another subject at randomly alternating points in time. The running total of earnings dynamically evolves according to random elements as well as the actions of counterparts. In these experiments subjects are neither informed of their position in the distribution of earnings nor of the realized distribution of opportunities to give or receive benefits. Our experiment manipulates this informational condition, to determine how, if at all, such information impacts cooperation and realized efficiency.

There is an experimental literature on how wealth inequality affects behavior in one-shot or finitely repeated social dilemmas, but the results appear mixed. Andreoni and Varian (1999) use the canonical trust game but provided varying show-up payments to subjects to induce inequality. They find no consistent effect of induced inequality. However, Greiner et al. (2012) find that initial inequality leads to greater trust in a repeated trust game with anonymous rematching because higher wealth is a clearer signal of previous untrustworthiness when initial conditions are equal.[1] Nishi et al. (2015), finds that players who are informed about partners' past behavior, in a networked public goods game, cooperate less when they are informed about the wealth of others. Gangadharan et al. (2015) find a negative impact of inequality on

[1]Interestingly, inequality models cannot describe behavior in two- and three-person trust games (Deck, 2001; Kagel and Wiley-Wolfe, 2001).

efficiency in a linear public good game where subjects can communicate with and reward others. In all of these studies, cooperating by investing own wealth to bestow benefits on others (i.e., social fungibility) is not part of a Nash equilibrium for a self-interested, rational player—unless one explicitly considers heterogeneity or introduces social components in preferences. By contrast, we study a game where full cooperation is a Nash equilibrium—even if players are homogeneous and self-interested—wealth is not socially fungible, and others' past conduct is kept private.

There is also mixed evidence on how externally-imposed payoff inequality affects the division of surplus in strategic settings. Goeree and Holt (2000) find that differences in fixed payments in laboratory bargaining games induce offers that are not consistent with Nash equilibrium but are consistent with a fair division of final payments. In a dynamic public goods experiment, Sadrieh and Verbon (2006) find no clear link between cooperation among partners and the exogenous distribution of rights to the surplus generated. Just as in Goeree and Holt (2000), differences in payoffs should not theoretically alter behavior in our design, but they do. Yet, unlike Sadrieh and Verbon (2006), we find evidence that the distribution of the running total of giving opportunities significantly affects cooperation in a dynamic setting.[2]

Finally, our study is situated in an experimental literature about how information influences the efficiency of outcomes. This literature is vast, as it straddles several research agendas, from the study of reputation (Bolton et al., 2005; Camera and Casari, 2009; Schwartz et al., 2000) , to the study of transparency and communication (Ellingsen and Östling, 2010; Huck et al., 2000; Isaac and Plott, 1981), from the impact of payoff asymmetries on coop-

[2]The evidence from non-strategic distributive choice experiments—where third-party spectators must select a division of resources—is also mixed (see Mollerstrom et al., 2015)

eration (Andreoni and Varian, 1999; Chen and Gazzale, 2004) to the study of the role of information in market and strategic experiments (Kagel and Levin, 1986; Nagel, 1995; Roth and Malouf, 1979). Although there are elements of commonality with all of these research themes, our most direct contribution is to the last strand of research. To discuss, we adopt a design where subjects cannot build a reputation in any treatment—they remain strangers in all interactions and can never establish reciprocal relationships. Second, all treatments are designed so that the past conduct of an individual always remains opaque—and players have no ability to communicate. Third, in all treatments cooperation symmetrically benefits players because there is always equal opportunity and players who face identical decisional situations have equal payoff matrices. When we manipulate the amount of information across treatments we find that less, not more, information is beneficial. This result is related to similar findings in market experiments and strategic bargaining games (see Smith, 1994, p. 119). Unlike those settings, however, we focus on information that is payoff-irrelevant, cannot disclose past conducts, nor can be used to build reputations. We find that the less informed players are about the distribution of past earning opportunities, the more cooperative they become. This pattern is not predicted by the standard application of folk theorem-type results. Yet, providing this information reduces cooperation: subjects who had many recipient opportunities tend to be more cooperative (and can be more forgiving) when no information about others is provided even if the efficient outcome is equally attainable in all informational settings.

3 Experimental design

In our experiment, a group of subjects played in an indefinite sequence of "helping games." Each game consists of a "donor" who is endowed with a good and a "recipient" who values the good more than the donor. The donor faces an individual decision problem: she can transfer the good to the recipient (*Help*), or she can consume the good (*Do nothing*). The recipient has no endowment and no action to take. All framing in the experiment was neutral. The structure of the game is in Table 1, while screenshots and instructions can be found in Appendix B.

	Donor	
	Help	*Do nothing*
Recipient	$g, 0$	$d-l, d$

Table 1: Payoffs in the helping game (g=25, d=6, l=2 points; 1 point=\$0.20).

If the donor helps, then the recipient earns g while the donor earns nothing. Otherwise, both subjects earn a default payoff, which is higher for the donor; d denotes the donor's default payoff and $d-l$ denotes the recipient's default payoff, with $g > 2d-l > 0$. There is *cooperation* in a meeting when the donor chooses to help the recipient; otherwise, there is *defection*. Given the payoffs, the donor's dominant action is to do nothing. Cooperation is *not* mutually beneficial but it is (socially) optimal as it maximizes surplus in the pair. In the experiment $d = 6$ points, $l = 2$ points, $g = 25$ points, where points is the experimental currency unit. The surplus from cooperation is $g-(2d-l) = 15$ points. The cost of cooperation to a donor is the difference in the two possible payoffs $0 - d = -6$ points; the benefit of cooperation to a recipient is her surplus $g - (d - l) = 21$ points; hence the benefit/cost ratio is 3.5.

Baseline session and supergame: A session involved 16 subjects in the lab at the same time, all exposed to the same treatment, which was chronologically divided into five distinct supergames. In a supergame, subjects interacted for an indefinite number of periods in fixed matching groups of size four. Hence, there were four concurrent supergames being played in a session. In each group, subjects were randomly re-matched into pairs at the start of each period, so there was a 1/3 probability of meeting the same person in two consecutive periods. Subjects did not know with whom they were paired nor did they know who was in their matching group in any supergame.

Every period, in each pair the computer randomly assigned the recipient role to one subject ("blue," in the experiment), and the donor role to the other ("red"), with equal probability. Hence, in every period half the subjects where recipients and half were donors. The random assignment of roles is a shock that affects the subject's earning potential for the period because recipients have a superior earning potential (25 points vs 6 points). This shock ensures equal economic opportunity going forward because future earning prospects are constant and identical for all participants. However, the random assignment of roles generates unequal economic results because it provides an exogenous source of variation in cumulative earnings. As the game progressed, some participants could be recipients more often than others, which would give them more chances of getting the higher payoff of 25 if the cooperative outcome was attained (and the higher payoff of 6, under full defection)

The duration of the supergame was uncertain because it was determined by a random continuation rule (as in Roth and Murnighan, 1978). A supergame began with 15 fixed periods after which successive periods occurred with probability $\beta = 0.75$. This continuation probability can be interpreted as the discount factor of a risk-neutral subject. *A priori*, the expected duration of a supergame was 18 periods because from period 15, in each period the supergame is expected to last 3 more periods. At the end of each period a computer drew an integer number between 1 and 100 with equal probability, which was then revealed to all subjects. A draw equal to or below 75 informed subjects that the supergame would continue (otherwise, it would end). [3]

At the end of each period, subjects observed whether or not the outcomes were identical in both pairs of their group. This form of anonymous public monitoring allows public detection of deviations from a social norm, and it could also simplify coordination tasks, but it does not allow agents to identify opponents (see screenshots from instructions in Appendix B). Hence, because individual histories remained private, subjects could neither build a reputation nor engage in relational contracting. Public monitoring ensured that the minimum discount factor supporting full cooperation in equilibrium was invariant across treatments (see next section).[4]

Supergames terminated simultaneously for all concurrent groups. After each of the first four supergames, subjects were placed into new four person matching groups and began playing another supergame. Matching groups were constructed so that no one was ever in a group with someone else more than once. Subjects were aware of this fact and, as a result, we have twenty unique groups per session. At the conclusion of a session, one supergame was selected randomly (Sherstyuk et al., 2013) and subjects were paid based upon their earnings in that supergame at the rate of $0.20 per point.

[3]This number could also serve as a public coordination device, at the group level.

[4]Subjects had access to information about past outcomes of every match in which they were involved. Each subject had pen and paper at their station.

Roles treatment: In this treatment donors observed explicit information about inequality before making their choice. At the start of any period after the first period in a supergame, one can measure the proportion of past periods in which a subject was a recipient (we call this the *recipient rate*). Unequal recipient rates give rise to inequality in past earnings, especially when cooperation rates are high given the greater spread in points.

In the Roles treatment, before making a choice, donors observed the normalized recipient rate for each group member. This information was called the "blue index" as it conveyed information about how often players had been in the blue role. The donor observed her blue index, the paired recipient's index, and the index of the two others in a random order. To facilitate comparisons, the average relative frequency of 0.50 was normalized to 100, so a value of $100 + x$ indicated a $x\%$ departure from the average.

Adding this index neither expands the action set relative to Baseline, nor affects payoffs in the stage game. The index expands the strategy set, as donors can condition their choice on the provided information, in periods $t \geq 2$. Otherwise, the treatment is identical to Baseline. In particular, index values (i) neither yielded points nor could be redeemed for points or dollars, and (ii) masked the identity of donors and preserved anonymity because they were not associated with individual identifiers and were unobservable to recipients.

Procedural Details: We recruited a total of 128 subjects through announcements at the University of Arkansas. All subjects recruited had no previous experience with this type of game.[5] After giving informed consent, subjects were seated at private terminals. Neither communication nor eye

[5]About 55% of subjects were males, and the rest female. The subject pool is composed of about 90% undergraduate students with the remainder being primarily graduate students although some faculty, staff, and non-university associated people are in the pool.

contact was possible among subjects at any time during the session. The experimenter publicly read the paper instructions at the start of the experiment, which were then left on the subjects' desks. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). On average, a session lasted 94 periods for a running time of approximately 120 minutes including instructions, a paid post-instruction comprehension quiz, and post-experiment payment. Average earnings were \$26.25 per subject ($\min = \$6.50, \max = \$55.50$) excluding a \$5 fixed participation payment and an average of \$2.10 ($\min = \$.75, \max = \$2.50$) from providing correct answers to the post-instruction comprehension quiz (\$0.25 for each of 10 questions). Only one randomly selected supergame from the session was paid. Table 2 provides summary details by treatment.

	Treatment	
Variable	*Baseline*	*Roles*
Blue Index	No	Yes
Group Size	4	4
Sessions	4	4
Subjects/Session	16	16
Supergames	80	80
Periods (avg.)	18.5	18.4
Salient Earnings		
average	\$26.38	\$25.94
min, max	\$8.75, \$54.00	\$6.50, \$55.50

Table 2: Sessions and treatments

4 Theoretical considerations

Here we show that in every treatment, groups can theoretically attain multiple Pareto-ranked equilibria, which range from full defection (no donor ever helps) to full cooperation (every donor always helps). Defection is the unique Nash equilibrium in a one-shot interaction because help is costly to a donor (0 instead of 6 points). It follows that *Full defection* is always an equilibrium because it consists of an indefinite repetition of the one-shot Nash equilibrium. Payoffs in the supergame are minimized under *Full defection* and are maximized only under *Full cooperation*, which is also an equilibrium because subjects could observe whether or not choices differed in their group, in each period. This form of anonymous public monitoring can be exploited to build a rule of cooperation supported by a punishment convention that is triggered if the rule gets broken. Specifically, full cooperation can be supported in equilibrium if a subject helps from the beginning of the game whenever she is a donor, but switches to defection forever after someone defects (Kandori, 1992, Proposition 1). When everyone adopts this strategy, then we say that cooperation is a *social norm*. Here, any defection is punished with permanent defection by the entire group.

Cooperation is an equilibrium when two conditions apply: in equilibrium, every donor prefers to help; out of equilibrium *no* donor prefers to help. The latter condition is immediately verified: once someone moves off equilibrium, that deviation is publicly observed. Hence, every donor defects thereafter and there is no longer an incentive to cooperate. The first condition requires checking that a donor cannot improve her payoff by moving off equilibrium (unimprovability criterion). In Appendix A we prove that this is the case as long as players are sufficiently patient, i.e., if $\beta \geq \beta^* := \dfrac{2d}{g+l}$.

Proposition 1. *In our experimental environment, full defection and full cooperation can be supported as an equilibrium.*

The threshold value β^* is the ratio between the cost of cooperation for a donor d and the surplus difference expected next period, amounting to $\frac{g+l}{2}$. The parameter β is the continuation probability of the game, 0.75. The condition $\beta \geq \beta^*$ is necessary and sufficient for the existence of a cooperative equilibrium. Based on the experimental design we have $\beta^* = 4/9$, so cooperation is an equilibrium in every treatment. Yet, there is no guarantee that full cooperation is realized instead of a lower-efficiency equilibrium.[6]

Three comments are in order. First, there is "equal opportunity" as players' future earning potential is identical, and independent of their past roles or behavior. Second, full cooperation supports income inequality as the realized sequences of donor and recipient roles inherently vary across subjects. This uncontrollable factor induces inequality in realized cumulative earnings but does not alter the structure of incentives because it does not affect continuation payoffs in the efficient equilibrium. Third, neither income inequality nor the underlying factor that generates it, can alter the power structure in the game because high-income participants have no greater control over the earnings of others than low-income participants. We make a further remark:

Proposition 2. *Revealing the distribution of past roles expands the strategy set but neither eliminates any of the equilibria that are possible in Baseline, nor increases the maximal equilibrium payoff relative to Baseline or alters the expected return from cooperation.*

The proof of this claim is simple. The Role treatment expands the strategy set compared to Baseline because donors can condition their help on the

[6]With public monitoring of defections, it is easily demonstrated that many other equilibria exist in all treatments, with efficiency degrees below that attainable under full cooperation and above that attainable under full defection.

information provided by the blue index. However, providing this information does not mean that subjects *must* use it. Subjects can always adopt strategies that ignore the index; therefore, the Role treatment does not eliminate any of the equilibria that are possible in Baseline. The second part of the statement follows from observing that the ability to condition behavior on the index is neither necessary nor sufficient to sustain full cooperation. It is not necessary because in all treatments the efficient outcome can be attained by conditioning choices on the actions known to have been taken in the group. It is not sufficient because the index masks the identity of counterparts, cannot be used to signal a cooperative intention, and does not reveal individual past conduct.

5 Results

We report five main results. The first three provide evidence that subjects conditioned their decisions on their own histories of roles in the supergame. The other results focus on how making salient the inequalities in past roles affected economic efficiency and cooperation patterns.

We start by observing that our Baseline experimental economies struggled to achieve the efficient outcome, even though it was theoretically within their reach.

Result 1. *In the Baseline treatment, cooperation generally increased over the course of a session, but did not reach 100%.*

Support for this result is provided in Tables 3 and 4. A subject's cooperation rate corresponds to the proportion of cooperative choices the subject took as a donor in the supergame. The first row in Table 3 shows that the mean cooperation rate in the Baseline treatment lies between 44 and 67 percent. As a comparison, the round-one cooperation averaged 63 percent.

Treatment	Supergame 1	2	3	4	5	*Total*	t=1	Realized efficiency ≤20%	≥80%	100%
Baseline	0.44	0.50	0.61	0.67	0.60	*0.56*	0.63	9	21	10
Roles	0.30	0.50	0.52	0.52	0.52	*0.47*	0.59	14	11	2

Notes: The unit of observation is a four-person group (N=16 per treatment, per supergame). Each cell reports the average proportion of cooperative choices in a supergame. The *t=1* column reports the average for period 1 of all supergames. The *Realized efficiency* columns report the number of groups that attained various cooperation levels (N=80 per treatment). No group achieved 0% cooperation.

Table 3: Average cooperation rate in a group

Table 4, reports marginal effects on the mean cooperation rate from a regression that pools data from all treatments. The regression includes controls for treatment effects, as well as a standard set of individual and other controls (e.g., subject's self-reported sex and duration of the supergame). To trace how experience with the task affects cooperation, we also include a dummy variable for each supergame above one, which is the base level. All *Supergame* coefficients are positive and significant, providing evidence that cooperation increased as subjects gained experience with the game. This evidence is in contrast with the dynamics of cooperation observed in social dilemmas with deterministic horizons (e.g., see Dal Bó, 2005; Palfrey, 1994) and is in accordance with what emerges from other studies on indefinite social dilemmas among strangers (Camera and Casari, 2009).

As our design admits multiple equilibria, it is possible that the moderate cooperation rate observed in Baseline is simply the result of different groups coordinating on different equilibria. As an example, if 45 out of 80 groups coordinate on the efficient outcome (100% cooperation) and the rest coordinate on full defection, then we obtain 56 percent average cooperation. The data do not support this hypothesis and, in fact, reveal considerable heterogeneity in individual cooperation rates.

Dep. variable: *cooperation rate*	Coeff.		S.E.
Roles	-0.110	***	(0.038)
Supergame 2	0.110	***	(0.026)
Supergame 3	0.179	***	(0.069)
Supergame 4	0.244	***	(0.042)
Supergame 5	0.235	***	(0.062)
Controls	Yes		
N	160		

Notes: Dependent variable: relative frequency of cooperation. Unit of observation: one group in one supergame (N=80 per treatment). Generalized Linear Model regression with robust standard errors (S.E.) adjusted for clustering at the session level. Controls include supergame duration, current and previous (set to 18 periods, in supergame 1), and the following individual characteristics: self-reported sex, and two measures of understanding of the instructions (response time and number of wrong answers in the quiz); the supergame duration and the response time regressors have negative and statistically significant coefficients. The regression includes interaction terms between treatment and supergame dummies (only *Roles* $\times$ *Supergame 2* is significantly different than zero with a positive coefficient and p-values<0.001). Symbols $***$, $**$, and $*$ indicate significance at the 1%, 5% and 10% level, respectively.

Table 4: Group cooperation: marginal effects.

Result 2. *In the Baseline treatment, (i) 14.0% of subjects never cooperated but no group coordinated on the inefficient equilibrium; (ii) 32.5% of subjects fully cooperated but only 12.5% of groups coordinated on the efficient equilibrium.*

Evidence for this is provided in Table 3 and Figure 1. In our design, realized efficiency corresponds to the group's average cooperation rate. Subjects did not coordinate on the inefficient equilibrium in the Baseline treatment: 7 groups out of 80 ended up below 20% efficiency, while no group attained 0% cooperation. On the other hand, 21 groups reached at least 80% efficiency and 10 of these attained the efficient outcome. This, and the increasing cooperation trend across supergames (Result 1), suggest that many subjects attempted to coordinate on the efficient equilibrium but were often unsuccessful due to heterogeneous behavior.

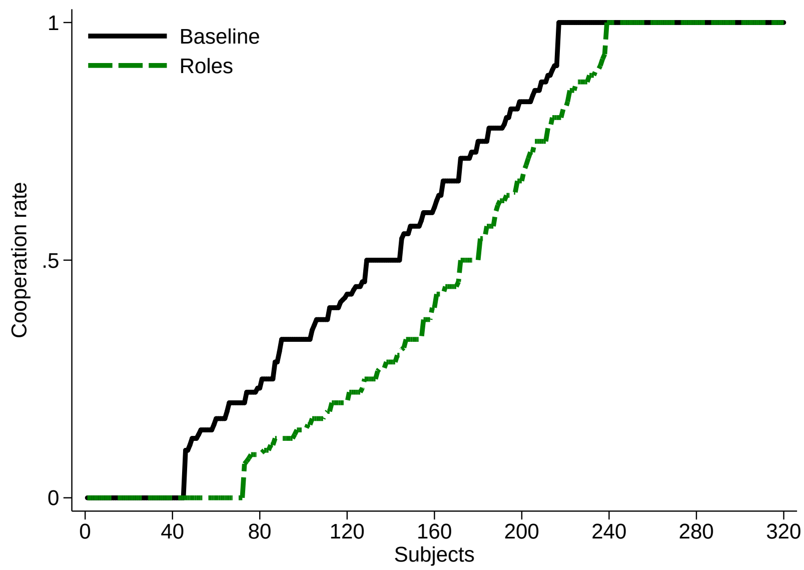

Figure 1: Cooperation rates: distribution (1 obs.=one subject in a supergame)

Figure 1, which reports the cumulative distribution of subjects' cooperation rates, provides evidence for this heterogeneity. In the Baseline treatment, quite a few subjects are classified by the strategy "always defect" (cooperation rate = 0 for 45 out of 320), twice as many by "always cooperate" (cooperation rate = 1 for 104 out of 320), and about half of subjects fall in between these extremes (171 out of 320). As only 10 groups attained full cooperation, clearly some subjects always cooperated even when some members of their group did not. A similar heterogeneity has been observed in other indefinitely repeated helping games (Camera and Casari, 2014; Camera et al., 2013).

Could disparities in the individual histories of shocks be a contributing factor? To investigate this possible behavioral factor, we consider the relative frequency of recipient roles for a subject in a supergame, which we call the *recipient rate*. The rate is 1 for a subject who was a recipient in all periods of a supergame, and it is 0 for a subject who was never a recipient. The distribution has mean and median equal to the expected value, 0.5.

Half of subjects lie between .4 and .6, but the rate varies between .18 and .81, meaning that in a supergame lasting 19 periods, some subjects were recipients in as few as three periods, while others were recipients in as many as fifteen periods. In a cooperative group, these differences would amount to an income gap of approximately 300 points ($375=25\times15$ versus 75 points) because variation in past roles is the only source of earnings variation. However, in Baseline the empirical correlation between a subject's donor frequency and income is -0.376 (p-value<0.001). We thus hypothesize that donors might have based their choices on their own histories of roles, with the intent to counteract unfavorable realizations of shocks. We find support for this hypothesis.

Result 3. *In the Baseline treatment, donors conditioned their choice to help on their own role history. Frequent recipients were more likely to cooperate than the average donor; occasional recipients were more likely to defect.*

Evidence is provided by Figure 2 and Table 5. We ran a probit regression where the dependent variable takes value 1 if the subject cooperated as a donor in a meeting (and is 0 otherwise). The regression controls for supergame effects, duration of previous supergame, period and individual fixed effects. To capture the trigger strategy discussed in Section 4, we include six dummy variables that determine the impact of a defection on the subsequent choice to switch to a punishment mode. The variable *Experienced Defection* takes value 1 if the subject did not always receive help in the past (and is 0 otherwise). As reported in Camera and Casari (2009, 2014), subjects might delay defection, thus the *n choice(s)* dummy variables are included. These variables take the value 1 if the donor made $n = 0, \ldots, 4$ choices after suffering the initial defection (and are 0 otherwise). The sum of *Experienced Defection* and each of the *n choice(s)* dummies traces the punishment strategy of the average subject.

Dep. variable:	Baseline			Roles		
=1 if donor helps	Coeff.		S.E.	Coeff.		S.E.
Supergame 2	0.027		(0.044)	0.177	***	(0.029)
Supergame 3	0.095		(0.066)	0.196	***	(0.035)
Supergame 4	0.112	***	(0.039)	0.167	***	(0.015)
Supergame 5	0.079	*	(0.048)	0.183	***	(0.026)
Roles history of donor						
Occasional recipient	-0.135	***	(0.018)	-0.066	***	(0.014)
Frequent recipient	0.126	***	(0.021)	-0.029		(0.052)
Punishment						
Experienced Defection	-0.355	***	(0.083)	-0.182	***	(0.045)
0 choice	0.127	***	(0.037)	0.092	***	(0.021)
1 choice	0.070	*	(0.042)	0.051	***	(0.014)
2 choices	0.033		(0.045)	0.060	*	(0.034)
3 choices	0.037		(0.048)	0.009		(0.038)
4 choices	0.055		(0.037)	-0.002		(0.022)
Controls	Yes			Yes		
N	2672			2680		

Notes: Dependent variable = 1 if donor helps, 0 otherwise. One observation=choice in a period (all periods>1). Controls include period fixed effects (dummy for five-period intervals 1-5, 6-10, 11-15, 16-20, 21-25, and > 25), duration of current and previous supergame (normalized to 18 for supergame 1), and individual characteristics (self-reported sex, and two measures of understanding of the instructions—response time and number of wrong answers in the quiz). Marginal effects are computed at the mean value of regressors of continuous variables. Robust standard errors with clustering at the session level.

Table 5: Probit regression on the choice to help: marginal effects

To determine if one's own frequency of past earning opportunities impacts behavior, we also include two dummy variables capturing the donor's role history up to that point in the supergame. For each period $t > 1$, let $r_t = 0, \ldots, t-1$ be the frequency of the subject's past recipient roles. We have $r_t = 0$ if the subject had yet to be a recipient in that supergame. We identify a donor's position in the distribution of r_t using three quantiles. The *Occasional recipient* regressor takes the value 1 for those in the bottom third of the distribution, and the *Frequent recipient* regressor takes the value 1 for donors in the top third. All remaining donors have an approximately balanced frequency of roles, and their category is taken to be the base in the regression.

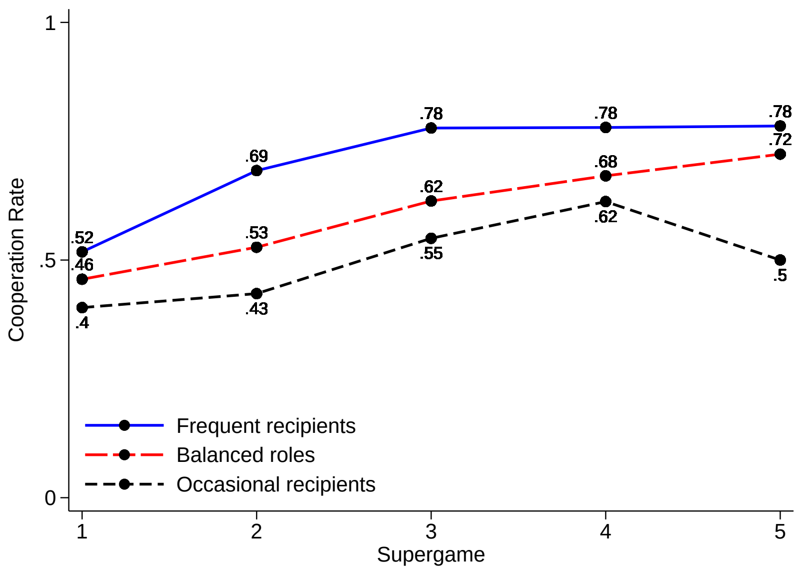

Figure 2: Cooperation in Baseline (1 obs.=one donor in a period)

Figure 2 reports cooperation rates of each quantile by supergame, and Table 5 reports the marginal effects on the probability of cooperating.[7] Consistent with the discussion of a social norm of cooperation in Section 4, the *Experienced Defection* regressor is negative and highly significant meaning that the average subject acted uncooperatively after experiencing a defection. However, in contrast with the use of a social norm to support cooperation, the *Frequent*

[7]For a continuous variable, the marginal effect measures the change in the likelihood to cooperate for an infinitesimal change of the independent variable. For a dummy variable, the marginal effect measures the change in the likelihood to cooperate for a discrete change of the dummy variable from its base level (0).

recipient coefficient is positive and highly significant. This means that the cooperation probability was greater for subjects who had very few opportunities to help others in the past compared to subjects who had an average frequency of past opportunities. Conversely, the cooperation probability significantly *fell* for subjects who were frequent donors compared to subjects who had a balanced frequency of past roles (the *Occasional recipient* regressor is negative and significant). Furthermore, the coefficients on the *Frequent recipient* and the *Occasional recipient* regressors are statistically different (p-value=0.000). This is evidence that donors varied their help in proportion to their own past roles, even if past shocks by design cannot influence the return from future cooperation.[8]

One may interpret Result 3 as stemming from subjects' desire to reduce disparities *in their group*, and not just to their own. We explore this idea in the Roles treatment where we make salient inequality in uncontrollable shocks because donors see the counterparts' relative frequency of past roles (their *blue index*). These disclosures do not theoretically alter the structure of incentives because they neither reveal the counterparts' past conduct, nor their future intentions. However, donors can vary their help in proportion to the counterpart's relative position.

We report that group cohesion suffered when extant differences in past earning opportunities were made salient.

Result 4. *Disclosing the relative positions in the distribution of past roles reduced cooperation and efficiency in the Roles treatment relative to the Baseline treatment.*

Support comes from Tables 3 and 4. The *Roles* dummy in Table 4 is

[8]The results in Table 5 are robust to including period fixed effects, with one dummy for each period. They are also robust to use of a panel logit regression—the panel variable being a subject in a session—with fixed-effects to control for unobserved subject-level heterogeneity.

negative and significantly different from zero (p-value=0.004). Efficiency is lower in the Roles treatment relative to the Baseline treatment because mean cooperation rates are lower in each supergame (see Table 3). Overall, the difference is 9 percentage points (0.56 vs. 0.47).

What lies behind this efficiency decline? Result 4 cannot be ascribed to differences in the allocation of roles in the two treatments. We can reject the hypothesis that the distribution of realized earning opportunities differed across any of the treatments we ran. A test for equality of distribution functions finds no statistically significant difference between the underlying distributions of subjects' recipient rates in a comparison between Baseline and Roles (Epps and Singleton test, p-value=0.852, N=320 per treatment).

A second hypothesis is that subjects acted uncooperatively out of a desire to reduce income inequality *in their group*, and not just to be opportunistic. Indeed, groups could eliminate almost all income inequality by coordinating on full defection, which minimizes per-capita income. Yet, no group did so in the Roles treatment, suggesting that reducing income inequality was not a primary goal. Groups could also maximize per-capita income by coordinating on the fully cooperative equilibrium. But this is also rarely the case and, in fact, coordination on the efficient outcome is less frequent than in Baseline; only 2 out of 80 groups managed to do so in the Roles treatment (see last column in Table 3). Furthermore, the Gini coefficient in income data from the Baseline treatment is 0.194, while it is 0.214 in the Roles treatment (unit of observation is one subject in a supergame, N=1280).[9]

[9]Income in the data also exhibits a higher degree of inequality than in counterfactual simulations were roles alternate as in the experiment but choices are exogenously fixed to either "do nothing" or "help." Considering the Roles treatment, the Gini for income drops to 0.026 in the counterfactual inefficient outcome where no-one helps, because recipients and donors' payoffs are similarly low (4 vs. 6); it increases to 0.132 in the counterfactual efficient outcome where donors always help because payoffs vary more (25 vs. 0); the Gini

A third hypothesis is that disclosing the relative positions in the distribution of past roles negatively altered the structure of incentives. Figure 1, which reports the cumulative distribution of subjects' cooperation rates by treatment, provides some supportive evidence. The distribution function in Baseline is first-order stochastically dominant; more subjects exceed any positive cooperation rate in Baseline than in the Roles treatment. A pairwise treatment comparison rejects the hypothesis that the distributions of subjects' cooperation rates are similar (Epps and Singleton test: p-value<0.001, N=320 per treatment). To uncover the source of the efficiency decline we examine individual behavior.

Result 5. *In the Roles treatment, frequent recipients were no more likely to cooperate than the average donor. Occasional recipients were more likely to defect.*

Figure 3, Table 5 and Table 6 provide support for this result. In contrast with Baseline, the *Frequent recipient* coefficient is negative and not significant (p-value=0.582), meaning that the cooperation probability did not increase for subjects who had many past opportunities to benefit from cooperation compared to the average donor. Moreover, we can also reject the hypothesis that the coefficients on the *Frequent Recipient* regressor are similar across treatments; This is achieved via a stacked regression (p-value=0.0101). We can also reject the hypothesis that the coefficients on the *Frequent recipient* and the *Occasional recipient* regressors are statistically different in the Roles treatment (p-value =0.409). The *Occasional recipient* regressor is negative, small and significant, meaning that the cooperation probability fell only slightly for occasional recipients compared to the average donor.[10]

coefficient for simulated data where choices are assumed random is 0.154.

[10]This coefficient is different across treatments (p-value=0.0359 from a stacked regression).

Overall, this is evidence that donors did not vary their help in proportion to their own past roles, as they did in the Baseline treatment. There is a decline in cooperativeness especially for those subjects who had a positive streak of past shocks. Figure 3 reveals that frequent recipients became less cooperative than the average donor, as the game progressed.

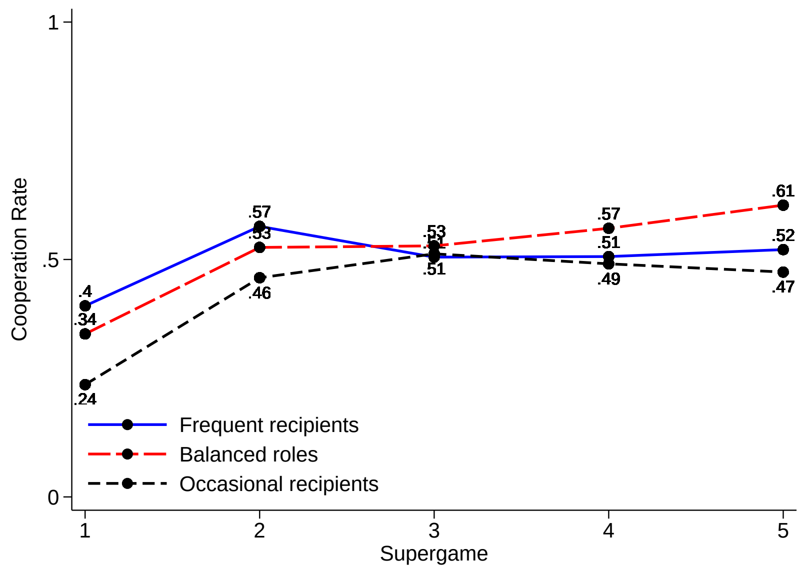

Figure 3: Cooperation in Roles (1 obs.=one donor in a period)

It is conceivable that this decline stems from donors cooperating less with counterparts known to be ahead in terms of past earning opportunities. To test this hypothesis we ran probit regressions where the dependent variable takes value 1 if cooperation occurred in a pair and is 0 otherwise. We classify pairs into four types depending on the roles history of recipient and donor, respectively. This is done using the *blue index.* The index can be calculated in any treatment, even when not shown to the subjects. A subject with a below-average blue index is “behind” (=B), and is “ahead” (=A) otherwise.

Based on this classification, we have four possible pairs (AA, AB, BA and BB), with corresponding dummy variables in the regression (BB=1 is the base level). For example, the regressor AB=1 if only the donor was a recipient in more than half of past periods. We also include the six trigger strategy dummy variables earlier described to soak up the effect of possible community punishment schemes.

Dep. variable:	Baseline		Roles	
=1 if donor helps	coeff.		coeff.	
Supergame 2	0.039		0.177	***
	(0.050)		(0.030)	
Supergame 3	0.114		0.197	***
	(0.082)		(0.034)	
Supergame 4	0.137	***	0.165	***
	(0.031)		(0.017)	
Supergame 5	0.091	*	0.185	***
	(0.049)		(0.028)	
Role history: donor & recipient				
BA	-0.011		-0.006	
	(0.013)		(0.037)	
AB	0.189	***	0.075	****
	(0.031)		(0.019)	
AA	0.143	***	0.066	***
	(0.047)		(0.009)	
Punishment dummies	Yes		Yes	
Controls	Yes		Yes	
N obs. (N subjects)	2672 (64)		2680 (64)	

Note: Probit regression. Dependent variable = 1 if donor helps, 0 otherwise. Base case = donor and recipient have a blue index < 100. B=blue index below average (=100), A=blue index equal to or above average. One observation=one choice in a period (all periods > 1). Punishment dummies include the same six variables in Table 5. Controls include period fixed effects (a dummy for periods in five-period intervals 1-5, 6-10, 11-15, 16-20, 21-25, and > 25), duration of current supergame and previous supergame (normalized to 18 for supergame 1), and individual characteristics (self-reported sex, and two measures of understanding of the instructions—response time and number of wrong answers in the quiz). Marginal effects are computed at the mean value of regressors of continuous variables. Robust standard errors with clustering at the session level.

Table 6: Donor's choice to cooperate: marginal effects.

Table 6 reports the marginal effects on the probability of observing cooperation in a pair. We have no evidence that donors based their help on the counterpart's blue index. The coefficients on the AB and AA regressors—which are both positive and significant—are not statistically different (Roles column, p-value=0.390).[11] There is also no evidence that donors acted less cooperatively with recipients ahead of them; the BA coefficient is close to zero and statistically insignificant (p-value=0.390).

Result 5 reinforces the view that subjects acted to reduce their own exposure to unfavorable past earning shocks. But why do we observe lower cooperativeness among those who had the most favorable shocks? A possible explanation is that those who were frequent recipients experienced a gambler's fallacy (Tversky and Kahneman, 1974) when confronted with how lucky they had been. Such an increased expectation of being in the donor role in the future would reduce their incentive to cooperate in the current period. An alternative explanation is that lucky subjects might reduce their own cooperation anticipating that unlucky subjects will become uncooperative.

6 Discussion

Our experiment finds support for the view that economic inequality affects cooperation and group cohesion. The result emerges from laboratory economies where inequality is theoretically neutral—it can neither alter the power structure in the group, nor the expected return from cooperation. In the experiment, participants benefit from developing norms of mutual cooperation over the long run, but helping carries a personal cost. A series of random

[11]As expected, they are also significantly smaller than the corresponding coefficients in Baseline. Pooled regression: p-value<0.000 for AB comparison, p-value=0.010 for AA comparison).

shocks ensures equal *future* earning potential for everyone, while inducing variation in *past* earning opportunities. In this setting, inequalities in past shocks should not alter the structure of incentives for payoff-maximizing players, hence should have no effect on cooperation. However, we find a clear behavioral influence.

First, there is evidence that participants conditioned their choices on their histories of shocks, cooperating in proportion to the frequency of their past earning opportunities. This type of backward-looking behavior is inconsistent with payoff-maximizing strategic behavior for two reasons. First, past shocks do not affect the returns from future cooperation, and so are theoretically payoff-irrelevant—by design past roles can neither influence the subject's future earning opportunities, nor her counterparts'. Second, the efficient outcome *cannot* be attained when subjects condition their choices on their own past roles. In the cooperative equilibrium, each donor must help unconditionally, independent of how many times she helped in the past. Off-equilibrium, past roles should not influence choices either since, under community punishment, a donor's best response is to defect.

A possible interpretation is that subjects were unwilling to follow a norm of mutual support when the associated benefit did not reflect *their* contribution to the prosperity of others. In our setup frequent donors have limited earning opportunities but abundant chances to increase others' fortunes. The opposite holds true for frequent recipients. Subjects might thus attempt to smooth these differences by conditioning their choices on their past roles as the game progresses. This type of backward-looking non-strategic behavior, which has not been documented before in indefinite social dilemmas, is consistent with experimental results from finite-horizon settings (Loch and Wu, 2008; Sonnemans at al., 1999). One could also interpret this behavior as consistent with the "luck egalitarianism" observed in non-strategic settings (Konow, 2000; Mollerstrom et al., 2015).

We find that group cohesion is elastic to certain types of inequality information. Disclosing the distribution of past earning opportunities—over which players had no control—crowded out group cooperation. This reaction supports the interpretation that subjects acted with the intent to counteract unfavorable realizations of earning shocks, which were made especially salient when subjects could observe the relative positions in the distribution of past roles. It is also possible that people are less willing to sanction uncooperative acts—switching to a defection mode—when the distribution of past shocks reveals that not everyone had an equal chance to benefit from cooperation.

References

Aghion, P. and J. Williamson. 1998. *Growth, Inequality, and Globalization.* New York: Cambridge University Press.

Anderson, Lisa R., Jennifer M. Mellor, and Jeffrey Milyo. 2004. Induced heterogeneity in trust experiments. *Experimental Economics.*

Andreoni, James and Hal Varian. 1999. Preplay contracting in the Prisoners' Dilemma. *Proceedings of the National Academy of Sciences of the United States of America* 96(19) 10933-10938.

Blonski M., P. Ockenfels and G. Spagnolo. 2011. Equilibrium Selection in the Repeated Prisoner's Dilemma: Axiomatic Approach and Experimental Evidence. *American Economic Journal: Microeconomics*, 3(3), 164-92.

Bolton, Gary E., Elena Katok, and Axel Ockenfels. 2005. Cooperation among strangers with limited information about reputation. *Journal of Public Economics* 89 (2005) 1457-1468.

Camera, G. and Casari, M. 2009. Cooperation among strangers under the shadow of the future. *The American Economic Review*, 99(3), 979-1005.

Camera, G. and Casari, M. 2014. The Coordination Value of Monetary Exchange: Experimental Evidence. *American Economic Journal: Microeconomics*, 6(1), 290-314.

Camera, G., Casari, M., and Bigoni, M. 2013. Money and trust among strangers. *Proceedings of the National Academy of Sciences of the United States of America*, 110(37), 14889-14893.

Chen, Yan and Robert Gazzale. 2004. When Does Learning in Games Generate Convergence to Nash Equilibria? The Role of Supermodularity in an Experimental Setting. *American Economic Review* 94(5), 1505-1535.

Dal Bó, P. 2005. Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. *American Economic Review* 95(5), 1591-1604.

Dal Bó, P. and Fréchette, G. 2011. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. *American Economic Review*, 101(1), 411-429.

Deck, C. 2001. A Test of Game Theoretic and Behavioral Models of Play in Exchange and Insurance Environments. *American Economic Review* 91 (5), 1546-1555.

Duffy, John and Jack Ochs. 2009. Cooperative Behavior and the Frequency of Social Interaction. *Games and Economic Behavior* 66 (2009), 785-812.

Ellingsen, Tore, and Robert Östling. 2010. When Does Communication Improve Coordination? *American Economic Review* 100, 1695-1724

Gangadharan, L., N. Nikiforakis and M. C. Villeval. 2015. Equality Concerns and the Limits of Self-Governance in Heterogeneous Populations. Manuscript.

Goeree, Jacob K., and C. A. Holt. 2000. Asymmetric inequality aversion and noisy behavior in alternating-offer bargaining games. *European Economic Review* 44(4), 1079-1089.

Greiner, B., Ockenfels, A., and Werner, P. 2012. The Dynamic Interplay of Inequality and Trust—An Experimental Study. *Journal of Economic Behavior and Organization* 81, 355-365.

Huck, S., H.T. Normann, and J. Oechssler . 2000. Does Information about Competitors' Actions Increase or Decrease Competition in Experi mental Oligopoly Markets? *International Journal of Industrial Organization* 18, 39-57.

Isaac, R. Mark, and Charles R. Plott. 1981. The Opportunity for Conspiracies in Restraint of Trade. *Journal of Economic Behavior and Organization* 2, 1-31.

Kagel, John H. and Dan Levin. 1986. The Winner's Curse and Public Information in Common Value Auctions. *American Economic Review*, 76(5), 894-920

Kagel, John H., and K. Willey-Wolfe. 2001. Tests of fairness models based on equity considerations in a three-person ultimatum game. *Experimental Economics* 4(3), 203-219.

Kandori, Michihiro. 1992. Social norms and community enforcement. *Review of Economic Studies*, 59, 63-80.

E. O. Kimbrough, V. Smith and B. J. Wilson, 2008. Historical Property Rights, Sociality, and the Emergence of Impersonal Exchange in Long-Distance Trade. *Am. Econ. Rev.*, 98(3), 1009-1039

Konow, James, 2000. Fair shares: accountability and cognitive dissonance in allocation decisions. *American Economic Review* 90, 1072-1091.

Loch, Christoph H. and Yaozhong Wu. 2008. Social Preferences and Supply Chain Performance: An Experimental Study. *Management Science* 54(11), 1835-1849,

Mollerstrom J., B. Reme, and Erik Sørensen 2015. Luck, choice and responsibility—An experimental study of fairness views. *Journal of Public Economics* 131, 33-40

Nagel, Rosemarie. 1995. Unraveling in Guessing Games: An Experimental Study. *American Economic Review*, 85(5), 1313-1326

Nishi, A. H. Shirado, D. G. Rand, and N. A. Christakis. 2015. Inequality and visibility of wealth in experimental social networks. *Nature*.

D.C. North, 1991. Institutions. *J. Econ. Persp.*, 5(1), 97-112.

New York Times. Inequality Troubles Americans Across Party Lines. By Noam Scheiber and Dalia Sussman, June 3, 2015

Palfrey, T.R. and Rosenthal, H., 1994. Repeated Play, Cooperation and Coordination: An Experimental Study. *The Review of Economic Studies*, 61(3), 545-565.

Pew Research Center. 2014. Emerging and Developing Economies Much More Optimistic than Rich Countries about the Future.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century.* Cambridge, MA: Harvard University Press.

Putnam, Robert. 2000. *Bowling Alone: The Collapse and Revival of American Community.*

Roth, Alvin E. and Michael W. Malouf. 1979. Game-theoretic models and the role of information in bargaining. *Psychological Review*, 86(6), 574-594.

Roth, Alvin E., and Keith Murnighan. 1978. Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical Psychology*, 17, 189-98

Sadrieh, A., and H.A. Verbon, 2006. Inequality, cooperation, and growth: an experimental study. *European Economic Review* 50, 1197-1222.

Schwartz, Steven T., Richard A. Young, and Kristina Zvinakis. 2000. Reputation without Repeat Interaction: A Role for Public Disclosures. *Review of Accounting Studies*, 5(4), 351-75.

Sherstyuk, Katerina, Nori Tarui, and Tatsuyoshi Saijo, 2013. Payment schemes in infinite-horizon experimental games. *Experimental Economics*16 (1), 125-153.

Smith, V. 1994. Economics in the Laboratory. *Journal of Economic Perspectives* 8(1), 113-131.

Sonnemans, Joep, Arthur Schram, Theo Offerman (1999) . Strategic behavior in public good games: when partners drift apart. Economics Letters 62, 35-41

Stiglitz, Joseph E. 2012. *The Price of Inequality: How Today's Divided Society Endangers Our Future.* W. W. Norton & Company, New York.

Tversky, Amos and Daniel Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science* 185 (4157): 1124-1131.

Appendix A

Proof of Proposition 1

This analysis is based on the existence of equilibrium proof in Camera et al. (2013). In each period $t = 0, 1, 2\ldots$ individuals in the group are matched in pairs, with uniform probability of selection. In each pair, a computer randomly determines who is the donor and who is the recipient (with equal probability). If cooperation (=Help) is the outcome, then g is the payoff to the recipient and for generality let a denote the payoff to the donor. If defection (=Do nothing) is the outcome, then d is the payoff to the donor and $d - l$ to the recipient. Period payoffs are geometrically discounted at rate $\beta \in (0, 1)$ starting from period $n > 0$.

The equilibrium payoff (=expected lifetime utility) at $t = 0$ is

$$v(n) := (n+1) \times \frac{g+a}{2} + \sum_{j=1}^{\infty} \beta^j \times \frac{g+a}{2} \;=\; \frac{g+a}{2} \times \left(n + \frac{1}{1-\beta}\right).$$

A player is a donor or a recipient with equal probability in each period, hence expects to earn $\frac{g+a}{2}$ in each period. The payoff $v(n)$ is increasing in n because payoffs are discounted by β in periods $t \geq n$.

The equilibrium payoff in the continuation game starting on any date $t \geq 0$, *before* any uncertainty is resolved, corresponds to

$$V_t = \begin{cases} v(n-t) & \text{if } \; t < n \\ v^* := \dfrac{g+a}{2(1-\beta)} & \text{if } \; t \geq n. \end{cases}$$

The equilibrium payoff of a donor at the start of any date t is

$$V_{dt} = \begin{cases} a + v(n-t-1) & \text{if } \; t < n \\ a + \beta v^* & \text{if } \; t \geq n. \end{cases}$$

We must check that in equilibrium donors have no incentive to defect; out of equilibrium, donors have no incentive to cooperate.

Defection is the dominant action off-equilibrium; i.e., it is always individually optimal to punish after a defection from equilibrium play is made public. To see this suppose a donor deviates by helping off equilibrium. She would earns a instead of d but her continuation payoff would not improve since everyone else keeps defecting—as prescribed by the rule of punishment. Since $d > a$, it is optimal to punish off equilibrium.

In equilibrium, cooperation is a best response in every period $t = 0, 1, \ldots,$ if $V_{dt} \geq \hat{V}_{dt}$. The left-hand-side denotes the payoff to a donor who cooperates; the right-hand-side denotes the donor's payoff when she moves off equilibrium under a one-time deviation. Such deviation is publicly observed, hence—when everyone follows the cooperative strategy—every donor will always defect in the future. The payoff to the deviator is thus

$$\hat{V}_{dt} = \begin{cases} \hat{v}(n-t) := d + (n-t)\dfrac{2d-l}{2} + \beta\dfrac{2d-l}{2(1-\beta)} & \text{if } \ 1 \leq t < n \\ \hat{v}^* := d + \beta\dfrac{2d-l}{2(1-\beta)} & \text{if } \ t \geq n \end{cases}$$

Now define

$$\Delta_t = V_{dt} - \hat{V}_{dt} = a - d + \frac{g + a - 2d + l}{2} \times \begin{cases} n - t + \dfrac{\beta}{1-\beta} & \text{if } \ t < n \\ \dfrac{\beta}{1-\beta} & \text{if } \ t \geq n \end{cases}$$

The minimum value of Δ_t is achieved for $t \geq n$. The implication is that cooperation is individually optimal in all periods t whenever

$$\beta \geq \beta^* := \frac{2(d-a)}{g+l-a}.$$