Are People Equally Other-Regarding When Picking a Partner vs Choosing an Allocation?*

Michele Belot† Marcel Fafchamps‡

July 2016

Abstract

We study how other-regarding behaviour vary across two decision contexts: when subjects make a pure allocation decision; and when they pick a partner. In both settings each subject's decision is final, as in a dictator game; and it affects their payoff and that of other subjects in the same way under both settings. We find that that subjects are less likely to sacrifice their own material wellbeing to increase that of others when selecting a partner in a large anonymous setting than when dividing a pie — even though the consequences on the material payoffs of others are identical. We interpret this differences as suggesting the application of different norms or heuristics: a pure allocation decision between four individuals including oneself resembles the decisions people make within a household, where norms of gift exchange and fairness apply; a partner selection decision resembles the decisions people make when competing for mates, where the pursuit of self-interest is acceptable.

JEL codes: A13; C91; D61; D63; D64

Keywords: Partnership formation, Other-regarding preferences, Markets, Inequality, Efficiency

*The authors thank the audiences of the IZA Conference on Network Economics in Bonn in June 2013, of the Workshop on Social Norms and Preferences in Lyon in October 2015 and thank Yann Bramoulle, Yves Zenou, Margherita Comola, Adam Szeidl and Philipp Kircher for useful comments. The project was funded by a grant of the John Fell Fund at the University of Oxford. An earlier version of this paper was circulated under the title "Experimental Evidence on Other-Regarding Preferences in Partner Selection".

†Contact author: School of Economics, 30 Buccleuch Place, University of Edinburgh, EH8 9JT Edinburgh, michele.belot@ed.ac.uk, +44 (0)131 651 5129

‡Freeman Spogli Institute for International Studies, Stanford University, Encina Hall 616 Serra Street Stanford, CA 94305-6055, fafchamp@stanford.edu

1 Motivation

There is considerable evidence that experimental subjects behave pro-socially in dictator games. This has been interpreted as suggesting that people have other-regarding preferences (e.g., Fehr and Schmidt 1999). There is also abundant experimental evidence that those who share less than fairly are subjected to second or third-party punishment – i.e., either by the recipient, as in the ultimatum game (e.g., Gueth et al., 1982), or by a third-party observer (e.g., Fehr and Gaechter, 2000). This suggests that other-regarding behavior may be influenced by social norms, and not just innate altruistic preferences.

In this paper, we ask how the *type* of decision people make affects the type of social norms that are triggered, and thereby, the type of other-regarding behavior exhibited. We focus on two types of decisions that relate to two key ways resources can be allocated in society: pure allocation decisions and partnership formation. Pure allocation decisions are an obvious way of allocating resources and have been studied extensively in the experimental literature, as mentioned above. Partnership formation, on the other hand, is another key mechanism by which resources are allocated. It is at the core of economic and social life: friendships, marriages, business partnerships or political alliances are a fundamental underpinning of economic and social exchange (McPherson, 2001). The sorting patterns resulting from partnership formation play a crucial role in the allocation of resources within a society, both in terms of efficiency and equity. Economists and sociologists have long been interested in how relationships are formed and who matches with whom, starting with the early work by Becker (1973).

We hypothesize that these two decision environments elicit different social norms and thus trigger different behavior. Pure allocation decisions resembles those people make within the family – e.g., dividing a pie between themselves and their spouse, children, or siblings. In such environment, fair division is the rule. In contrast, partner selection decisions put individuals in competition with each other – often across households. In such decisions, people often behave in a rival manner. Yet partner selection choices often have allocative consequences on others. We are interested in finding out whether these consequences are taken into account when selecting a partner.

This is best illustrated with a couple of examples. Our first example has two brothers. One is handsome and rich; the other is ugly and poor. Both compete for the favors of the same beautiful girl. Is it reasonable to expect the handsome brother to bow out of the competition because 'it is more fair that his brother gets the girl' since he is ugly and poor?[1] We suspect not.

Our second example involves a game of doubles tennis. There are four players, two weak and two strong. They all want to win. How do we team them up? One possibility is to let players select their partner individually. In that case, we expect each strong player to select the other as partner, leaving the two weak players to team up with each other. Alternatively, let one person select between two team configurations: (weak-weak, strong-strong) or (weak-strong, weak-strong). Chances are this person will select the (weak-strong, weak-strong) configuration because it is a fairer, more egalitarian allocation. What is interesting about this example is that any strong player can deliver the fair configuration simply by not selecting – or refusing to team up with – the other strong player. In other words, the two decision environments are identical in terms of the final allocations players can impose. But since they evoke different norms of behavior, they may result in different outcomes – and hence different payoffs. If this is indeed the case, it will demonstrate that it is futile to seek to identify purely intrinsic other-regarding preferences without reference to the decision context.

Our experimental design is organized in a manner similar to the doubles tennis example. We compare other-regarding preferences in two experimental settings: when subjects make a pure allocation decision; and when they pick a partner. In both settings each subject's decision is final, as in a dictator game; and it affects their payoff and that of three other subjects. The final payoffs implied by the subject's choices are the same under both settings, and these payoffs are shown to experimental subjects in the same manner under both settings. But the decision that subjects make is different. We test whether subjects are less likely to sacrifice their own material wellbeing to increase that of others when selecting a partner than when dividing a pie — even when the

[1] The example works equally well if two sisters are competing for the same man.

consequences on material payoffs are identical.

This paper studies the role of other-regarding preferences in the partner selection process. The literature on decentralized matching is agnostic as to what role these preferences may play. But there is now a well established literature on social preferences (Fehr and Schmidt 1999, Bolton and Ockenfels 2006, Charness and Rabin 2002, Cooper and Kagel 2013) showing that people are willing to transfer resources to others, in particular if these transfers reduce inequality, and that the role of other-regarding preferences seems particularly salient in small groups and among kin (e.g., Fehr and Schmidt 1999, Ledyard, 1995, Camerer, 2003, Henrich et al. 2005).

In contrast, competition can deliver an efficient outcome even if all participants act to maximize their own material welfare and nothing else. For markets, the original insight goes back to Adam Smith's shoemaker parable, and it has been verified in numerous market experiments (e.g., Smith 1962). Becker (1973) offers a similar result in the context of competition for mates. In a review of the experimental evidence, Bowles (1998) observes that the more the experimental situation approximates a competitive (and complete contracts) market with many anonymous buyers and sellers, the less other-regarding behavior is observed.[2] Vernon Smith (1998) points out that these two aspects of human nature are not contradictory but apply to different contexts. Selfish behavior maximizes the gains from impersonal market exchange, while cooperative behavior maximizes the gains from non-market personal exchange.

So far, little is known as to what role other-regarding preferences play in partner selection. The literature on decentralized matching describes match formation as a market-like process while the literature on other-regarding preferences suggests that such preferences are particularly strong in small partnerships. So the question is: do people apply competitive heuristics when searching for a partner (i.e. behave selfishly); or do they behave pro-socially, as they do once these partnerships are formed?

Abstracting from strategic considerations, we want to examine if other-regarding behavior is

[2] This finding fits with the two apparently opposite views of Adam Smith who argues in the *Wealth of Nations* (1776) that self-interest prevails in markets, while acknowledging the pro-sociality of human nature in the *Theory of Moral Sentiments* (1759).

similar in a situation where people are asked to choose a partner among many relative to a situation where they asked to make an allocation decision in a small group. Going from one setting to the other involves two sources of variation. One is the type of decision: partnership formation versus allocation decision; the second is the size of the group people are in. We hypothesize that both features make it harder to realize the implication of one's choice on others. That is, it may be cognitively more difficult to understand the implications of one's decision on others when one is asked to choose a partner among many, relative to a simple dictator-like allocation decision.

We conduct an experiment in the laboratory. We consider an environment with two "categories" of people intended to represent, in an uncontextualized way, the two sides of a partner selection choice. The two categories we consider are labeled "Addition" and "Multiplication" and correspond to an effort task performed at the beginning of the experiment. We invite 24 people in each experimental session. Participants are assigned at random to one of these two categories and, based on their performance in the real effort task, to one of two types – 'high' and 'low'. In each session, we have 6 participants of each type and each category.

The first treatment we introduce is a *dictator allocation decision* involving themselves and three other subjects. Each group of four subjects includes two high and two low types from each category. Participants are asked to choose between two divisions of payoffs among the four of them, with no reference to partner selection. In a second treatment, we present the same allocation problem, but the decision is framed as a *choice between a high or low partner*. In the third treatment the allocation problem is also framed as a choice of partner, but unlike in treatment 2 the decision is among the twelve players of the other category. In treatments 2 and 3, experimental subjects are shown the payoff distributions associated with all four types of partnerships. In all three cases, the decision of one of four players is selected at random and implemented in a unilateral manner to determine the payoff of all four. This allows us to abstract from strategic considerations when we infer preferences from choices.

Treatment 3 is the partner selection context that we are most interested in. Going from Treatment 3 to Treatment 2 changes the size of the group (large group vs small group) but keeps the decision as a partner selection decision. Going from Treatment 2 to Treatment 1 changes the type of decision from partner selection to allocation decision. In the first treatment, the implications that decisions have on others' material payoffs are obvious. Although they are the same in Treatments 2 and 3, they become increasingly less salient as we move from Treatment 1 to Treatment 3.

We consider different types of other-regarding preferences: pro-social preferences (Fehr and Schmidt 1995; Bolton and Ockenfels 2006); invidious preferences (Blanchflower and Oswald 2004; Fafchamps and Shilpi 2008); and preferences for efficiency (Charness and Rabin 2002). In our design we separately vary the implications and saliency of partner choice on efficiency and income distribution. We study unilateral partnership formation - i.e., partnership formation without mutual consent – in order to identify preferences separately from strategic considerations in partnership formation.[3]

In our three experimental treatments we find that participants mostly follow their material self-interest. This is in line with numerous studies of behavior in market games. We find a significant difference between large groups (Treatment 3) and small groups (Treatment 2), but a much smaller difference between Treatments 2 and 1. This suggests that subjects are more likely to display other-regarding preferences whenever they clearly perceive that their decision has a direct impact on others. In Treatment 3 they are less likely to sacrifice their individual payoff to increase aggregate efficiency. As a result, aggregate efficiency falls since the design rules out the operation of competitive forces. There is little or no statistical difference between Treatments 1 and 2.

Because we experimentally assign participants to different average payoffs, we can test whether deviations from pure selfish behavior vary with expected income from the experiment. We find that high payoff agents are on average more altruistic than low payoff agents, but significantly less so in a partner selection environment. Similarly, low payoff agents are less reluctant to reduce others' payoff in a partner selection environment, particularly if the number of participants is large.

[3] Comola and Fafchamps (2015) run a partner selection experiment with competition. They show that experimental subjects almost always converge to a stable match. The implication is that subjects are quite adept at competing for mates in a partner selection game. Here the focus is instead on the possible role of other-regarding preferences when selecting a mate.

The paper is structured as follows. Section 2 presents the experimental design and the different treatments. The testing strategy and empirical results are presented in Section 3 and we conclude in Section 4.

2 Experimental Design

Participants play in sessions of 24[4] and participants in a session only play one treatment – i.e., we use a 'between subjects' design. We refer to the Appendix 1 for a detailed description of the Experimental Protocol.

2.1 Stages

Within each session the experiment is divided into two stages. In the first stage, the pool of participants is divided equally and randomly into two categories A and M.[5] The two categories are intended to represent, in an uncontextualized way, the two sides of a matching game, e.g., bride-groom, employer-employee, or hospital-intern. Having been assigned to a category, each participant individually completes a computerized task. Subjects are asked to do simple calculations for a period of 3 minutes – additions or multiplications depending on the category, A or M, to which they have been randomly allocated. Based on their performance in the task relative to other participants in the same session and category, they are assigned one of two types – bottom 50% or top 50%. Here we denote the two types simply as 'high' and 'low'.

In the second stage of the experiment, participants play six rounds of an allocation game. No feedback is provided to participants until the end of the experiment, at which time they are only told their final payoff. Since payoffs are based on one randomly selected round, it is impossible for participants to infer the choices of other participants. The purpose of this is to rule out repeated games and strategic play: each round is de facto a dictator game with anonymous others and no feedback.

[4] In 3 sessions (one for each treatment), the number of people who reported for the experiment was less than 24. Hence the experiment was played in groups of 20 instead. This does not affect the experiment except in treatment 3.

[5] In the experiment these categories are refereed to as 'Addition' and 'Multiplication' – see below.

In the first two treatments, participants are assigned to groups of four. Each group consists of a low and high type participant from each category. In treatment 1 (T1) choices are presented as a selection between two 'pies' divided into four possibly unequal slices – see Instructions in Appendix 4. For each round, each participant in each group selects his favorite pie/allocation. At the end of the session, one subject from each group and one round are selected at random, and his/her choice determines the payoff of all four participants in the group. The structure of the game thus resembles a dictator game with four players.

In treatment 2 (T2) participants are asked to indicate whether they would prefer to form a partnership with the low or high type participant from the other category. They are shown how the payoffs would be distributed for each possible partnership (high A/high M, high A/low M, low A/low M, low A/high M). They are also made aware of the implications of their choice for the 2 other participants in their group (see Instructions for Treatment 2 in Appendix 4). For example, if a participant in category M, say, selects the high type in category A as partner, this also determines the payoff of the two remaining players since they can now only be matched with each other. As in treatment 1, final payoffs are determined by randomly selecting one round and one subject from each group, and letting his/her choice of partner determine the payoff of all four participants in the group.

In treatment 3 (T3), all 24 participants in a session play a partner selection game together as follows. Each participant is asked to select one of two possible types of partners, high or low, among players of the other category (see Instructions for Treatment 3 in Appendix 4). Since the 24 participants are divided equally between categories A and M, and subsequently divided equally between low and high type within each category, there are six participants in each category $\times$ type. Because players are anonymous and there is no feedback, the choice of each player resembles treatment 2 except for the larger number of players. They are also shown how the payoffs would be distributed for each possible partnership, as in treatment 2. But they are not told anything explicitly about the implications of their choice for others. To understand these implications, they need to understand that by choosing a partner of a certain type, they prevent one person from their own category to be matched to a partner of that particular type.

Payoffs in treatment 3 are determined as follows. We first randomly select one category (A or M) at random as well as one of six rounds. We then aggregate the choices of the participants in the selected category. If there is excess demand for one type, then the scarce type is allocated in a random manner between those who have expressed a preference for it. For instance, say category A is selected. Of the 12 participants in category A, 8 have selected to match with a high type. Since there are only 6 high types in category M, each of the 8 participants is allocated to a high type match with a probability equal to 6/8, and a low type match with probability 2/8. The 4 participants who have selected to match with a low type get their choice of type. The payoff determination process is explained in detail to participants before the experiment (see Instructions in Appendix 4 for details).

In treatment 1, it is obvious by design that selecting one pie affects one's payoff and that of three other players. In treatment 2, it is clear to each player that their choice affects the payoff of the partner they choose. Given the payoff determination rule, they can also deduce that selecting one partner de facto forces the other two players together. Alternatively they may follow a market logic and convince themselves that their choice directly affects only one other player, and hence that their other-regarding preferences only apply to that player.

In treatment 3 players can, as in treatment 2, clearly see the effect of their choice on their – yet to be determined – *partner*. But by taking one possible partner away from the choice set of other participants, they also de facto limit the choices of other players. This effect is less salient than in treatment 3, however.[6] Because of the various rounds of randomization that take place before payoffs are assigned, treatment 3 may blur the sense of responsibility that participants associate with their actions. We call this the 'dilution hypothesis'. A formal presentation of this hypothesis is given Appendix 2.

[6]See Appendix 3 for a detailed discussion.

2.2 Payoffs

Payoffs in the second stage of treatments 2 and 3 represent how gains from matching are shared between matched partners. Gains are always shared equally if both partners are of the same type (both high or both low). If partners are of different types, the division of payoffs differs from game to game so as to vary the efficiency and equity of an heterogeneous match. We call each payoff matrix a scenario.

In a session all participants play 6 different scenarios. The scenarios are the same for all the participants in the same session. The same set of scenarios/payoff matrices is used for the three treatments. We use a set of 17 different scenarios (see Table 2). In 12 scenarios, payoffs are such that an heterogeneous match – and hence negative assorting (e.g., low A with high M) – is more efficient; in 6 scenarios, payoffs are such that positive sorting is efficient. We also have one scenario where positive and negative sorting are both efficient.

We also vary how the gains from a match are divided between the two participants. The three different division rules used are summarized in Table 1.

Table 1. Possible sharing rules

	Low A	High A
Low M	1/2,1/2	1/2,1/2 1/3, 2/3 1/6, 5/6
High M	1/2,1/2 2/3, 1/3 5/6, 1/6	1/2,1/2

The scenarios are summarized in Table 2 where are reported the payoffs associated with the different types of partnerships (or different allocations in treatment 1). The different scenarios have been chosen so that we can assign as unambiguously as possible a sequence of choices made by an individual subject to a specific preference archetype, provided that the subject's choices are all consistent with that archetype. We come back to this in Section 3.2.

High types have higher payoffs on average, and low types low payoffs. To help identify other-regarding preferences,[7] we introduce a further payoff differentiation between low A and low M payoffs such that low M types get, on average, lower payoffs than low A types. Throughout the analysis we present results broken down by these three payoff categories.

It is important to realize that in each T2 and T3 scenario participants ultimately choose between two possible pairings: (1) high A-high M / low A-low M and (2) high A-low M / low A - high M. The first corresponds to positive assorting, the second to negative assorting. Each of the pairings has an associated total payoff for the four participants affected by someone's choice. Hence to each scenario is associated two possible efficiency values. In some scenarios positive assorting is efficient; in others negative assorting is efficient. Treatment T1 mimics these differences albeit without explicit assorting. In all but 3.8% of the observations one form of assorting is more efficient than the other. However, whether positive or negative assorting maximizes someone's material payoff varies by category and type and according to the sharing rule associated with a particular scenario (see Table 1).

Treatments may affect expressed preferences in different ways. Because treatment 1 resembles the allocation process that takes place within a household, where some members (e.g., the person in charge of grocery shopping) make choices for other members, it may elicit preferences that are more other-regarding. One possibility is that participants pay more attention to the efficiency cost of their action. Another is that they seek a more equal allocation of payoff, or that they are less spiteful.

Because treatment 3 has features resembling an anonymous market environment, participants may use market heuristics when making choices, feeling empowered to pursue their own material welfare and thereby ignoring the consequences of their choice on others. Treatments 2 and 3 also raise the possibility that players may prefer to match with someone of their type (e.g., high or low) – see for instance Currarini and Mengel (2011) for recent evidence of homophily in experiments with randomly assigned types.

[7] Specifically, we need this design to distinguish between selfish preferences and maximin preferences for low types.

3 Analysis and results

The experiment took place at the experimental laboratory of the Centre for Experimental Social Sciences in May and November 2011.[8] We ran 12 sessions (4 sessions per treatment). 308 participants took part in total. Participants earned £10.28 on average.

The first objective of the experiment is to test whether play is systematically different between treatments and, in particular, whether decisions are more sensitive to (1) the payoffs of others and (2) to efficiency considerations as the saliency of the implications of choices increases. The second objective is to test what type of preference is most likely to account for observed behavior. We report on these two objectives in turn.

3.1 Sensitivity to the payoffs of others

Throughout we denote by π_i the payoff of subject i and N_i be the set of four players affected by i' choice.[9] We define efficiency as the sum of the payoffs of the four participants affected by the choice of a single player.

3.1.1 Variation across treatments

We begin by showing that play varies systematically with treatments. The right panel of Table 2 presents the raw aggregate distribution of choices for each scenario. We present the share of participants choosing a partner of the same type. We observe large differences across types and across scenarios.

In the upper left panel of Table 3 we report, for each treatment, the proportion of times a subject picks the choice that maximizes his or her individual payoff. The first three columns present the average for each of the three treatments, while the last columns reports the test statistic and p-value for a χ^2 square test of equality of means across treatments. Results are presented separately for subjects assigned to high payoff, low payoff in the A category, and low payoff in the M category.

[8] We ran the sessions corresponding to Treatment 3 in June 2011 and those corresponding to Treatments 1 and 2 in November 2011.

[9] In treatment 3, the exact individuals affected are unknown prior to ex post randomization, but their payoffs are known since, by design, they belong to the four category × type groups. See Appendix 3 for a formal demonstration.

As indicated earlier, low payoffs for subjects experimentally assigned to the A category are higher than for those assigned to the M category. We see that most choices are those that maximize the subject's material payoff. There is little difference across treatments, and whatever differences are present are not statistically significant. We also note that the proportion of choices consistent with selfish preferences is highest for low average payoff participants and lowest for high average payoff participants.

The upper right panel of Table 3 shows the proportion of times that a subject picks the choice that maximizes the payoff of the other three players affected by their choice. As before, we report proportions under each of the three treatments as well as the results of a χ^2 square test of equality of means across treatments. Results are presented separately for high payoff subjects, and low payoff subjects assigned to the A or M category.

The proportion of choices that maximizes others' payoffs decreases systematically as we go from treatment T1 to treatment T3. This is true for all payoff categories, but is strongest and highly significant for high payoff participants. For the lowest payoff types, those in category M, the difference is not statistically significant at standard levels. For the low payoff participants in the A category, the difference is significant at the 10% level. This suggests that participants take others' payoff less into consideration as we move from treatment T1 to treatments T2 and T3, but more so if their average payoff is higher. We also note that the magnitude of the increase in proportions is similar between treatments T1 and T2 and between T2 and T3.

In many cases, the choice that maximizes the payoff of others also maximizes one's own payoff. This explains why it is possible for, say, high payoff subjects in treatment T3 to maximize their own payoff in 80.8% of the cases while at the same time maximizing the other players' payoff in 52.7% of the cases. To investigate this issue further, in the lower half of Table 3 we split the choices that maximize others' payoff depending on whether doing so also maximizes one's own payoff or not. A 'good Samaritan' spirit corresponds to the case where participants maximize others' payoffs even though doing so reduces their own. This case is displayed in the lower right panel of Table 3. In the lower left panel we show the proportion of choices that maximize others' payoffs when doing so also maximizes the subject's own material payoff. Such choices are consistent with efficiency consideration but are less remarkable from an equity point of view.

In the lower left-hand panel we find that participants are more sensitive to aggregate efficiency in treatment T1 than in partner-selection treatments T2 and T3. This is true for all three payoff categories of participants, but only significantly so for high payoff participants. From the lower right-hand panel, we see that the "good Samaritan" spirit is highest in treatment T1 and lowest in T3. The difference is large in magnitude and significant at the 1% level for high and category A low payoff participants; for category M low payoff participants the difference is smaller in magnitude and only significant at the 16% level. These low payoff participants thus seem to focus exclusively on their own payoff, and show little regard for the payoff of others. One possible explanation is that they are trying to make up for having been assigned to the lowest payoff category at the onset of the experiment.

To further verify our results, we investigate whether treatments affect the sensitivity of participants' choice to differences in payoffs between the two options they face. Let $\Delta\pi_i \equiv \pi_i^h - \pi_i^l$ be the gain in i's payoff from choosing a high partner (or equivalent allocation in T1). Similarly let $\Delta\pi_{-i} \equiv \sum_{j \neq i, j \in N_i} (\pi_j^{h_i} - \pi_j^{l_i})$ be the gain in the payoff of the other three players affected by i's choice. To calculate the marginal effects of $\Delta\pi_i$ and $\Delta\pi_{-i}$ on the probability of choosing a high partner (or equivalent allocation in T1) we estimate a regression model of the form:

$$y_i = \Delta\pi_i \otimes T \otimes X + \Delta\pi_{-i} \otimes T \otimes X + \varepsilon_i \tag{1}$$

where $\Delta\pi_i \otimes T \otimes X$ is shorthand for all the possible interaction terms between them and similarly for $\Delta\pi_{-i} \otimes T \otimes X$.[10] Regression (1) is estimated using a linear probability model with standard errors clustered at the session level.

We report in Table 4 the estimated marginal effects with their t-value for each of the $T \otimes X$ combinations. What the Table reveals is that choice sensitivity to the payoff difference between the two options is comparable across treatments: a one unit increase in payoff gain increases the

[10] In other words, we include terms in π_i, T, X, TX, $\pi_i T$, $\pi_i X$ and $\pi_i TX$ where T and X are themselves vectors of dummies. Similarly for π_{-i}, T and X.

probability of choosing the more beneficial option by 5 to 6 percentage points across all subject types and treatments. Since the standard deviation of $\Delta\pi_i$ is 4.67, this is a large effect. There is a decrease in sensitivity to own payoff in treatment T1 for high types and low A types, but the difference is relatively small in magnitude, not statistically significant, and we observe nothing similar for low M types.

In contrast, choice sensitivity to $\Delta\pi_{-i}$ varies dramatically across treatments and subject types. For high types – who on average earn higher payoffs – sensitivity to $\Delta\pi_{-i}$ is absent in treatment T3 but present in the other two treatments. The effect is large in magnitude: a one standard deviation (i.e., 7.67) increase in $\Delta\pi_{-i}$ raises the probability of choosing high by 7.5% in treatment T2 and 11.8% in treatment T1. In treatment T3 the effect is numerically 0. This suggests that among these players considerations of altruism or efficiency are eliminated in T3, a finding that is consistent with the dilution hypothesis.

Results are different for low types. Here we find that, under anonymous partner selection T3, low type participants are at the margin *less* likely to choose a high partner if other players benefit more from that choice, controlling for their own payoff gain. The effect is large in magnitude, especially for low A types: a one standard deviation increase in $\Delta\pi_{-i}$ reduces the probability of choosing high by 11.4% for low A and 7.1% for low M types.

These findings are to be read in the context of the literature on inequality aversion (e.g., Fehr and Schmidt 1999, Okada and Riedl 2005). Experimental evidence has suggested that individuals display a desire to reduce the difference between their payoff and that of others both from above and from below. In other words, if a subject has a high payoff relative to other participants, this subject is often observed taking redistributive actions that reduce the difference between her payoff and that of other participants. This is consistent with the behavior of high types in our experiment, who have higher average payoffs and, in treatments T2 and T1, are more likely to choose an action that increases the payoff of other participants who, on average, earn a lower payoff. In other words, high types often choose an action that reduces the difference between their payoff and the lower payoff of others.

Also according to inequality aversion, a subject who has a low payoff relative to others often takes actions that increase her payoff at the expense of others. This is what we observe low types do in treatment 3: controlling for their own payoff, low types – who on average earn lower payoffs – are more likely to take an action that reduces the payoff of others. What is interesting is that, in our experiment, the two behaviors do not coexist: altruism (inequality aversion from above) is only present in treatments that emphasize the effect one's choice has on several others; envy or spite (inequality aversion from below) is only present in the treatment that blurs the effect of one's choice on others. This suggests that the dilution effect reduces altruism/efficiency considerations, but not envy which is, rather, exacerbated by anonymity. In other words, the two sides of inequality aversion respond differentially to an anonymous market/partner selection environment: altruism is blunted by it, while envy is increased. In our experiment, the two effects together combine to reduce the efficiency of participants' choices.

3.1.2 Choices and efficiency

We next investigate the relationship between treatment and the efficiency of individual choices. We ignore the 60 observations in which choices generate the same total payoff for the four players affected by ego's choice. We focus on whether the choice ego makes is efficient or not. In treatment T1 participants choose the efficient allocation in 70% of the observations. This proportion falls to 67% in T2 and 61% in T3. This difference is statistically significant at the 1% level.

We reproduce this finding in the first column of Table 5. Treatment T3 is the default category so that reported coefficients capture the efficiency gain of individual choices in T2 and T1 relative to T3. Furthermore, if we test average efficiency between T1 and T2, the difference is not statistically significant. In other words, participants are more likely to opt for an efficient allocation when the effect they have on others' payoff is more obvious. In contrast, individual choices are significantly less efficient in an anonymous partner selection setting.

In the other columns of Table 5 we disaggregate the results and regress the efficient choice dummy on treatment for each of the three payoff categories separately. The results indicate that the

T1 and T2 are associated with more efficient choices for both high and middle payoff participants. For low payoff participants, the treatment effect is not significant although, in terms of average efficiency across all treatments, low payoff participants are not statistically different from other payoff categories. It appears that an anonymous partner selection setting enables higher payoff participants to behave in a more selfish manner while in treatments T1 and T2 they may feel moral pressure to behave altruistically towards lower payoff participants. In contrast, low payoff participants do not seem to have this concern, perhaps because they feel more entitled to pursue their self-interest, having been assigned to the lowest payoff category at the onset of the experiment.

3.2 Preference archetypes

In this subsection we adopt a more structural approach and assign participants to archetypes summarizing the form of their other-regarding preferences. We test whether this assignment varies with treatment. The idea behind the approach is that people behave in ways that may vary with the decision context, and that their behavior can be approximated by preference archetypes. If we know what archetypes best capture the behavior of a large fraction of the population in a given environment, we may be in a better position to predict what types of behavior to expect in that environment.

Scenarios were designed to facilitate the assignment of participants to the six different archetypes listed below (Charness and Rabin 2002, Cooper and Kagel 2013). These archetypes were chosen because they are easily identified within the framework of our experiment.

1. *Selfish:* Chooses the allocation that maximizes π_i.
2. *Efficient:* Chooses the allocation maximizes $\overline{\pi}_i = \sum_{j \in N_i} \pi_j$.
3. *Equity only*: Chooses the allocation that minimize absolute inequality defined as $\sum_{j \in N_i} |\pi_j - \overline{\pi}_i|$
4. *Spiteful*: Chooses the allocation that maximizes one's relative payoff $\pi_i - \frac{1}{3} \sum_{j \neq i, j \in N_i} \pi_j$
5. *Maximin*: Chooses the allocation that maximizes the minimum payoff among the four affected individuals $\max \min \pi_j$, $j \in N_i$.

6. *Homophily*: Chooses a partner of the same type (high or low)

We seek to assign each participant to the archetype that best describes their behavior in the six rounds. Different participants may follow different archetypes. We start by identifying participants who follow a single archetype perfectly over all six rounds, and we observe what proportion of participants we can assign in this manner. For high payoff types, payoffs are such that assignment is typically unambiguous. For low types, the choices made by a given subject may be consistent with more than one archetype, depending on the set of six scenarios they faced. Next, we then introduce the possibility of deviations from the behavior predicted by a single archetype. At the end of the section, we allow for the possibility of hybrid preferences.

In Table 6 we present the result of our first calculation. For each archetype k, we calculate $\Delta u_i^k \equiv u_k(\pi_i^h) - u_k(\pi_i^l)$ where preference function $u_k(.)$ is that corresponding to archetype k. For each subject we then count the proportion of rounds (out of a maximum of 6) for which the subject behaves in accordance to archetype k, that is, for which $y_i = 1$ if $\Delta u_i^k > 0$ and $y_i = 0$ if $\Delta u_i^k < 0$. We ignore cases in which $\Delta u_i^k = 0$ because they are uninformative. We say subject i makes choices consistent with archetype k if this proportion is 100%. Depending on scenarios and player type, choices made over six rounds may be consistent with more than one archetype.

Results indicate that the archetype most consistent with observed choices is the 'selfish' archetype in which people only care about their material payoff. We do, however, observe systematic differences across treatments and types. Fewer participants are assigned to the selfish archetype in treatment T1 than in other treatments. The difference is strongest for high types, confirming earlier results that suggest these participants behave in a less selfish manner when the experiment is framed as an allocation process rather than as partner selection. Low types tend to behave more in accordance to the selfish archetype than high types, again confirming earlier results.

Turning to other archetypes, we note that very few participants are assigned to the 'efficient' archetype, that is, have a utility function that would lead them to always choose the most efficient allocation. We do, however, note that the efficient archetype is more common in treatment T1, and among high types. This is consistent with some of our earlier findings that suggest efficiency concerns among some participants, but it reminds us that 'efficient' is not a suitable description of the way participants behave on average.

A small proportion of participants behave in agreement with the equity only archetype. This archetype assumes that participants always choose the allocation that results in the smallest level of payoff inequality. This archetype is found mostly among low types for whom this overwhelmingly coincides with self-interest (100% overlap in choices between the two archetypes for low A types and 91% overlap for low M types).

A larger proportion of participants fit the spiteful archetype, that is, make choices that minimize the difference between their payoff and that of others. This archetype is found most often among participants to treatment T3, in line with earlier results emphasizing that the anonymous partner selection treatment seems to encourage spiteful choices relative to the other two treatments. Caution is however warranted because 89% of the low M types whose decisions fit the spiteful archetype also fit the selfish archetype.

A non-negligible proportion of low A types follows maximin preferences, i.e., they make choices that maximize the minimum payoff to any player. Since low A types nearly always are those players getting the minimum payoff, the data shows that maximin always coincides with self-interest for low A types. Overlap is also common (90%) among low M types. Hence this finding should not be given too much weight. Finally, contrary to other experiments that have documented that participants naturally assort according to randomly assigned type (e.g., Currarini and Mengel 2012), we find very little evidence of homophily in our data: very few participants systematically choose a partner of their type, i.e., low with low or high with high.

At the bottom of the table we report the proportion of participants who were assigned to multiple archetypes, and those whose behavior does not fit any. The last panel of Table 5 confirms that, by experimental design, high types can nearly always be unambiguously assigned to a single archetype. We were unable to do the same for low types because, with six rounds and four players, there were not enough degrees of freedom in the experiment to design scenarios that are sufficiently informative for both high and low types. This too is confirmed in Table 5. Multiple assignment is most common among low A types, a point that has already been touched upon in the earlier presentation.

The last row of Table 6 indicates that a sizeable proportion of participants do not fit any of the six archetypes we considered. Lack of fit is most noticeable in treatment T1, and this is true for all payoff categories. The behavior of participants in this treatment is less well explained by the simple preference models we considered. But in all treatments there is a sizeable minority of participants whose behavior is not consistent with any of the archetypes we considered.

One possible explanation for this is that people make mistakes: they may have preferences that follow one of our archetypes but, due to inattention or lack of interest, they sometimes make choices that do not correspond to their underlying preferences. To investigate this possibility, we estimate a mixed maximum likelihood model. The starting point of the estimation methodology is the observation that:

$$\Pr(y_i = 1|\pi_i, \pi_{j\neq i}, T_i) = \sum_{k=1}^{K} \Pr(y_i = 1|\pi_i, \pi_{j\neq i}, T_i, k) \Pr(u = u_k|T_i) \quad (2)$$

where π_i and $\pi_{j\neq i}$ denote the four payoffs potentially entering the preference utility u_k of archetype $k = \{1, ...K\}$. Since π_i and $\pi_{j\neq i}$ are randomly assigned in the experiment, we can ignore correlations between payoffs and preferences. But we allow preferences to differ across treatments, hence the conditioning on T_i. A similar probability can be derived for $y_i = 0$. Since, for a given treatment T_i, $\Pr(u = u_k|T_i)$ is a constant, we denote it as γ_{kT}.

Next we assume that the probability of choosing action $y_i = 1$ increases in Δu_i^k, the utility gain from choosing $y_i = 1$ that is associated with payoffs π_i and $\pi_{j\neq i}$ when preferences are given by $u_k(.)$. To formalize this idea, we borrow from Luce (1959) and write:

$$\Pr(y_i = 1|\pi_i, \pi_{j\neq i}, T_i, k) = \frac{e^{\sigma_T \Delta u_i^k}}{1 + e^{\sigma_T \Delta u_i^k}} \quad (3)$$

Parameter σ_T captures how sensitive decisions are to Δu_i^k in treatment T. If $\sigma_T = 0$, the choice between $y_i = 0$ and $y_i = 1$ is random and does not depend on payoffs. If σ_T is arbitrarily large, expression (3) tends to 1 if $\Delta u_i^k > 0$ and to 0 if $\Delta u_i^k < 0$ – which corresponds to the case where choices are perfectly predictable once we know someone's archetype and the payoffs they face. Intermediate values of σ_T capture situations in which participants systematically diverge from random play in the direction predicted by archetype k. The model assumes that participants are more likely to take the decision predicted by their archetype the larger the utility gain Δu_i^k is between the two choices. In other words, participants make more mistakes when the difference in payoff is small. Mixed models of this kind have successfully been fitted to experimental data (e.g., Andersen et al. 2008, Null 2012).

The likelihood function has the form:

$$L(\gamma_k, \sigma | y_i, \{\Delta u_i^k\}, T_i) = \sum_{k=1}^{K} \gamma_{kT} \left(\frac{e^{\sigma_T \Delta u_i^k}}{1 + e^{\sigma_T \Delta u_i^k}} y_i + \frac{1}{1 + e^{\sigma_T \Delta u_i^k}} (1 - y_i) \right) \tag{4}$$

and is estimated separately for each treatment, ensuring that $0 < \gamma_{kT} < 1$ and imposing that $\sum_{k=1}^{K} \gamma_{kT} = 1$.[11] Once σ_T and $\{\gamma_{1T}, ...\gamma_{KT}\}$ have been estimated, we compute, for each subject, the posterior probability that their choices follows a particular archetype.[12] Accumulating across rounds for each individual i, we get the posterior probability that i follows archetype k:

$$\Pr(k|\{y_i\}) = \frac{\Pr(k) \Pr(\{y_i\}|k)}{\Pr(\{y_i\})} \tag{5}$$

where $\{y_i\} = \{y_i^1, y_i^2, ...y_i^6\}$ is the set of decisions made by i over the six rounds. Since $\widehat{\gamma}_{kT}$ and $\widehat{\sigma}_T$ vary across treatment, $\Pr(k|\{y\})$ also varies across treatment for the same set of choices made. Once we have $\Pr(k|\{y\})$ for each subject, we look at how accurate the predictions are for different individuals, i.e., how accurately they are estimated to follow a given archetype.

[11] Estimation is achieved by numerical optimization in Stata. To ensure that all Δu_i^k have the same weight in the estimation, we normalize them to all have a unit standard deviation. Convergence difficulties can arise when $\gamma_{kT} \approx 0$ for some k, or when the Δu_i^k are too correlated across archetypes (akin to multicollinearity).

[12] Let $\Pr(k|y_i = 1)$ denote the probability that individual i is of archetype k if he/she sets $y = 1$. The starting point of our calculation is the following relationship that holds for each choice i makes:

$$\begin{aligned} \Pr(k|y = 1) &= \frac{\Pr(k) \Pr(y = 1|k)}{\Pr(y = 1)} \\ &= \frac{\Pr(k) \Pr(y = 1|k)}{\sum_k \Pr(k) \Pr(y = 1|k)} \end{aligned}$$

For simplicity of exposition, we have omitted the dependence on π and T. Unconditional probability $\Pr(k)$ is estimated by $\widehat{\gamma}_{kT}$ while $\Pr(y = 1|k)$ is obtained from expression (3) using estimated $\widehat{\sigma}_T$.

Estimates of posterior probabilities are summarized in Table 7. The first panel of the Table reports the average of estimated posterior probabilities (5) calculated using parameters $\widehat{\gamma}_{kT}$ and $\widehat{\sigma}_T$ estimated using (4). The second panel of the Table assigns each individual to an archetype if their $\Pr(k|\{y_i\})$ exceeds 0.5 for one k – which can happen at most for one k, but could happen for none. As it turns out, all our participants are assigned to one archetype, which suggests that the method was able to infer everyone's type with accuracy.[13] We do, however, note that the estimated value of $\widehat{\sigma}_T$ is smallest for treatment T3 ($\widehat{\sigma}_{T_3} = 6.60$ with a t-value of 5.59), intermediate for T2 ($\widehat{\sigma}_{T_2} = 14.48$ with a t-value of 4.25) and largest for T1 ($\widehat{\sigma}_{T_1} = 255.29$ with a poorly estimated t-value). This suggests that participants fit one of the six archetypes better in T3 than in T1. This confirms that we are less able to predict individual behavior in T1, perhaps because participants' choices reflect preferences other than our six archetypes – more about this later.

We find that most participants fit the selfish archetype best. This is true for all treatments and all player types. Secondly, high types are systematically more likely to act selfish in partner selection treatments, particularly T3. This confirms earlier results. Third, participants with the lowest average payoff (low M types) overwhelmingly play selfish, even in the pie allocation treatment. This too is consistent with earlier observations. Fourth, the proportion of players classified as maximizing efficiency increases as we move from treatment T3 to T1. This is particularly noticeable for high types, mirroring earlier observations. But only a small proportion of players are best described as following this archetype.

Fifth, a number of players are classified as having spiteful preferences, that is, as choosing strategies that reduce the difference between their payoff and that of others even if it means reducing their own payoff. (If they had chosen to maximize their own payoff, they would have been classified as selfish, not spiteful.) The proportion of spiteful is highest in treatment T3, especially among low types. This findings is also in agreement with earlier results.

Finally, we find that a number of low types, especially those in the intermediate payoff category

[13] In fairness, we should point out that we do not correct predicted archetypes for sampling error in parameter estimates. But given how strong assignment is, introducing the correction, if it were possible, would not greatly modify our results.

A, are classified as maximin players. This happens more frequently in treatment T1, suggesting that players in that treatment have some equity concerns, but these are not adequately captured by the equity only archetype – which only fits a small number of participants overall, and none of them in treatment T1. Equal pie sharing thus does not appear to be the top priority for any of the participants.

As robustness check, we also investigate a hybrid model that combines own payoff with an other-regarding preference u_i^m and estimate regressions of the form:

$$\Pr(y_i = 1) = \alpha_0 + \alpha_1 \Delta\pi_i + \alpha_i(\Delta\pi_i)^2 + \beta_1 \Delta u_i^m + \beta_2(\Delta u_i^m)^2 + \beta_3 \Delta\pi_i \Delta u_i^m + \varepsilon_i \tag{6}$$

where $\Delta\pi_i$ and Δu_i^m represent the material payoff and utility gain associated with choice $y_i = 1$, respectively. As before, all Δu_i^m are normalized to have the same unit standard deviation. Model (6) is estimated with preference archetypes (2) to (6) above as well as with altruistic preferences and with inequality aversion.[14]

For archetypes, the results, not shown here to save space, confirm earlier findings: efficiency considerations are significant in T2 and especially T1; invidious preferences are significant in treatment T3; maximin and equity-only preferences appear with the wrong sign, especially in treatment T1, a finding that is consistent with spiteful preferences; and there is little evidence of homophily. Results regarding altruistic preferences are contrasted. In treatment T3 altruism has the wrong sign, a result consistent with spiteful preferences. But in treatments T2 and especially T1, Δu_i^m becomes positively significant while $\Delta\pi_i$ no longer is. This suggests that altruist preferences do a reasonably good job of predicting participants' choices in T2 and T1 – but not in T3. In contrast, inequality aversion Δu_i^m is either statistically non-significant, or appears with the wrong sign. How much weight we should ascribe to the latter findings is unclear, given that the experiment was not

[14] For altruist preferences, we set $u_i \equiv \ln(\pi_i) + 0.5\ln(\pi_{-i})$ where $\pi_{-i} \equiv \sum_{j \neq i, j \in N_i} \pi_j$. Parameter value 0.5 captures an intermediate level of altruism. For inequality aversion we follow Okada and Reidl (2005) and posit a utility function of the form

$$u_i^{10} \equiv \ln(\pi_i) - \frac{1}{2} \sum_{j \neq i, j \in N_i} |\ln(\pi_i) - \ln(\pi_j)|^- - \frac{1}{3} \sum_{j \neq i, j \in N_i} |\ln(\pi_i) - \ln(\pi_j)|^+$$

and vary N_i to include either only i's partner or the three individuals affected by i's choice.

designed to test altruism or inequality aversion directly, and so has little power.[15]

4 Discussion and conclusions

We have reported on an experiment designed to test whether people exhibit different other-regarding preferences depending on whether the choices they make are framed as an allocation problem or a partner selection problem and depending on the size of the group from which they select a partner.

We find that when choices are framed as a partner selection problem instead of a pure allocation problem, agents are less likely to sacrifice their own material well-being to increase the well-being of others – but more willing to sacrifice a higher payoff to reduce the difference between their payoff and that of others. Similarly, they behave more selfishly if they are in a larger group setting.

These findings are broadly consistent with the literature on inequality aversion (e.g., Fehr and Schmidt 1999, Okada and Riedl 2005), but with a twist. Experimental subjects with a higher than average payoff exhibit some altruism or concern for efficiency, but more so in treatments are couched as an allocation problem or a choice of partner in a small group. When asked to select a partner in a large anonymous setting, high payoff players no longer display any sign of altruism and simply maximize their own material payoff. In contrast, subjects with the lowest average payoff display no altruism or concern for efficiency in the small group setting, but exhibit spiteful preferences in a large anonymous setting. In other words, we get the two 'sides' of inequality aversion (altruism and spite), but not in the same setting.

These findings raise a number of issues. Fafchamps (2011) argues that economic development requires a change in allocation processes away from allocation within the household or extended family to allocation within markets or within hierarchical organizations. This transformation requires a change in social norms from risk sharing in long-term gift exchange to contract compliance in an anonymous market setting. In a gift exchange allocation process, efficiency requires that individuals make choices that are altruistic or efficiency-seeking. In market exchange, efficiency can

[15] The coefficient of correlation between $\Delta\pi_i$ and Δu_i^m exceeds 0.9 for altruist preferences and each inequality aversion measure we tried.

be achieved through competition alone; altruistic or efficiency-seeking behavior is not required. To the extent that the behavior of our experimental subjects can be interpreted as reflecting context-specific norms, they fit this pattern to a large extent. We do, however, also find that less fortunate participants occasionally select a partner so as to prevent them from achieving a higher payoff. It is unclear whether competition is sufficient to counter the inefficiency produced by such choices. More research is needed.

The findings also raise a more fundamental question: Why, even with minimal contextualization, do human subjects respond the way they do to differences in the frame in which choices are made? Is it possible that the human brain processes moral choices in a way that systematically reduces altruism and reinforces spite in an anonymous partner-selection setting relative to a small group setting?

The literature on trolley experiments (e.g., Greene 2012) suggests one possible avenue of enquiry so far ignored by economists, namely that people feel less guilty about the consequences of their actions when these consequences seemingly depend on mechanical devices, random events, and choices made by others.[16] If correct, this interpretation suggests that markets – and partner selection problems in large populations – blunt other-regarding preferences by diluting the perceived effect that actions have on the welfare of others, thereby eliciting less guilt for failing to follow norms of acceptable behavior that apply in small groups. We offer in Appendix 2 a simple model of such preferences.[17] Further work is needed on the origin of other-regarding preferences, and especially the extent to which they are shaped by the decision environments over which altruistic norms apply.

[16] Mikhael (2011) goes so far as to suggest that this is because the human brain processes moral choices of cause and effect by applying syntaxic rules. This means that 'pushing the man to his death' generates more guilt than 'pushing the button that activates the lever that opens the door that pushes the man to his death' even though the ultimate consequence is the same.

[17] The difference between treatments T2 and 3 similarly is in the *process* by which payoff are generated, not in the actual payoffs themselves. Dilution as defined here can thus be seen as a situation in which people's preferences depend not just on payoffs but also on the way these payoffs are obtained. If this interpretation is correct, this paper can be seen as following in the footsteps of Charness and Rabin (2003) who demonstrated that people have preferences over process, not just final outcomes.

5 References

References

[1] Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström, (2008). "Eliciting Risk and Time Preferences," *Econometrica*, 76(3): 583-618, 05.

[2] Becker, G. (1973), "A Theory of Marriage: Part I", *Journal of Political Economy,* 81(4): 813-46.

[3] Blanchflower, D.G. and A.J. Oswald (2004), "Well-Being over Time in Britain and the USA", *Journal of Public Economics*, 88(7-8), July 2004

[4] Bolton, G.E. and A. Ockenfels (2000), "ERC: A Theory of Equity, Reciprocity, and Competition", *American Economic Review*, 90(1): 166-193, March.

[5] Bowles, S. (1998). "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions", *Journal of Economic Literature*, 36(1): 75-111.

[6] Charness, G. and M. Rabin (2002). "Understanding Social Preferences With Simple Tests", *Quarterly Journal of Economics,* 117(3): 817-869,

[7] Camerer, C., Thaler, R. H. (1995). "Anomalies: Ultimatums, dictators, and manners", *Journal of Economic Perspectives*, 9(2): 209-219.

[8] Camerer, C. (2003), *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press, Princeton NJ.

[9] Comola, Margherita and M. Fafchamps (2015), "An Experimental Study of Decentralized Link Formation with Competition", Stanford University and Paris School of Economics (mimeograph)

[10] Currarini, S. and F. Mengel (2012), Identity, Homophily and In-Group Bias, *FEEM Working Paper* No. 37.2012

[11] Fafchamps, M. and F. Shilpi (2008). "Subjective Welfare, Isolation, and Relative Consumption", *Journal of Development Economics*, 86(1): 43-60, April

[12] Fafchamps, M (2011). "Development, Social Norms, and Assignment to Task", *Proceedings of the National Academy of Sciences (PNAS)*, 108(Supplement 4): 21308-15, December

[13] Fehr, E. and K.M. Schmidt (1999). "A theory of fairness, competition, and cooperation", *Quarterly Journal of Economics*, 114: 817–868.

[14] Fehr, E. and S. Gaechter (2000). "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90(4), 980-994.

[15] GÃ¼th, W., R. Schmittberger and B. Schwarze (1982). "An experimental analysis of ultimatum bargaining", *Journal of Economic Behavior and Organization*, 3(4), 367-388

[16] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. Smith, Hill, K. Gil-White,F., Gurven, M. , Marlowe,F. W., Patton, J.Q., Tracer, D. (2005), "Economic man in cross-cultural perspective: Behavioral experiments in 15 small-scale societies", *Behavioral and Brain Sciences*, 28(6): 795-815.

[17] Hoffman, E., K. McCabe and V. Smith (1996), "Social distance and other-regarding behavior in dictator games", *American Economic Review,* 86: 653-60

[18] Greene, Joshua (2010). "Of Trolleys and Cheaters: Automatic and Controlled Influence on Moral Judgment", video available at http://www.santafe.edu/research/videos/play/?id=a768d1f4-06e5-4533-8ebe-9ba3d3a5c131,

[19] Greene, Joshua (2012). *For The Greater Good: How the Moral Brain Works and How it Can Work Better*, Penguin Press, New York, (forthcoming)

[20] Ledyard, John O. (1995). "Public Goods: A Survey of Experimental Research", in J. Kagel and A. E. Roth, *Handbook of Experimental Economics*, 111-194. Princeton: Princeton University Press.

[21] Luce, R.D., (1959). *Individual choice behavior*, John Wiley., Oxford, England

[22] Luce, R. Duncan (2010). "Behavioral Assumptions for a Class of Utility Theories: A Program of Experiments", *Journal of Risk and Uncertainty*, 41: 19-27.

[23] McPherson, M., L. Smith-Lovin and J. M. Cook (2001). "Birds of a Feather: Homophily in Social Networks", *Annual Review of Sociology*, 27: 415–44.

[24] Mikhail, John (2011). *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*, Cambridge University Press, Cambridge

[25] Null, Claire (2012). "Warm glow, information, and inefficient charitable giving", *Journal of Public Economics*, (forthcoming)

[26] Okada, Akira and Riedl, Arno, (2005). "Inefficiency and social exclusion in a coalition formation game: experimental evidence," *Games and Economic Behavior*, 50(2): 278-311, February.

[27] Smith, Adam. (1759) 1976. *The theory of moral sentiments*, Liberty classics, edited by D. D. Raphael and A. L. Mactie. Indianapolis:Liberty Press.

[28] Smith, Adam. (1776) 1909. *Wealth of nations*, Harvard Classics. volume 10, edited by C. J. Bullock. New York: P. F Collier

[29] Smith,V.L. (1962). "An experimental study of competitive market behavior", *Journal of Political Economy*, 70(2): 111-37.

[30] Smith, V. (1998), "The Two Faces of Adam Smith", *Southern Economic Journal*, 65(1): 1-19

[31] Young, P. and M.A. Burke (2001), "Competition and Custom in Economic Contracts: A Case Study of Illinois Agriculture", *American Economic Review*, 91(3): 559-573

[32] Cooper, David J. and John H. Kagel (2013). "Other-Regarding Preferences: A Selective Survey of Experimental Results", *Handbook of Experimental Economics*, Vol 2, J. H Kagel and A. E Roth (eds)

6 Appendix 1. Experimental protocol

We consider three different versions of the second stage. The first is the "*partnership formation in large groups*" version, whereby participants are organized in groups of 24 divided equally into A and M categories. participants are asked to choose a partner type from the other category . If their choice is implemented, we randomly match them with one of the 12 people from the other category. The second version is the "*partnership formation in small groups*" version, whereby participants are put in groups of 4 (a high and a low type in the A category and a high and a low type in the M category). participants are asked to choose whether they would like to form a partnership with the high or low type in the other category. The last version is the "*earnings division in small groups*" version whereby participants are not asked to choose a partner, but are simply asked to choose between two divisions of earnings between four players. The scenarios we propose in these three treatments are perfectly equivalent in terms of their implications for efficiency and income distribution. The only variation is the salience of these implications. Importantly, we provide the same information in all treatments.

Treatment T3 - Partnership formation - large groups

Here each participant is asked to report whether they would prefer forming a partnership with a low or high type from the other category. The scenarios are presented on an answer sheet (see Appendix 4). They are shown the distribution of earnings associated with the four different types of partnership (high A - high M, low A - low M, high A - low M, and low A - high M), including those that do not involve them. These choices are illustrated graphically with colored pies. The size of the pie represents the total earnings to be shared. Each earnings division corresponding to each partnership is represented with a pie division. Participants are asked to tick one of two boxes (top 50% or bottom 50%) at the bottom of the answer sheet.

The partnership formation is implemented as follows. One of the categories (A or M) is chosen randomly ex-post to be the leader in partnership formation (which means that the partnerships will be implemented according to their reported preferences, irrespectively of the reported preferences made by participants in the other category). If there is more demand for one type than is available, then the scarce type is allocated in a random manner between those who have expressed a preference for it.

Table 2 describes the different parameter configurations corresponding to the different situations. Note that the payoffs are not always symmetric for low types. To be able to distinguish between selfish preferences and maximin payoff, we designed a situation where the payoff of the low type individual is not the minimum payoff. In those situations, the gains from partnership are not shared equally between the A and M partners. The addition partner gets a larger share. In those cases, a low A type who wishes to maximize the minimum payoff would choose a high M type, while if she wishes to maximize her own payoff, she would choose a low M type.

Treatment T2 - Partnership formation - small groups

In the second treatment, participants are told that they are randomly allocated in groups of 4, composed of 2 people who did the multiplication task (one high, one low) and 2 people who did the addition task (one high, one low). They are asked again whether they would prefer forming a partnership with a low or high type from the other category. The scenarios are presented on an answer sheet (see appendix 4) in a manner similar to treatment 1, except that we now write explicitly the implication of the partnership decision for the other people in the group. They are asked to tick one of two boxes (top 50% or bottom 50%) at the bottom of the answer sheet (see Appendix 4)

Treatment T1 - Earnings division

In treatment participants are told that they are randomly allocated in groups of 4, composed of 2 M participants (one high and one low) and 2 A participants (one high and one low). They are asked to choose between two distributions of earnings that correspond to the earnings distribution in T3 and T2 (see Appendix 4). Each earnings distribution is represented by a pie division. The main difference with T3 and T2 is that to each choice is associated a single pie that represents the division of earnings between the *four* people in the group, instead of two pies for each of the two partnerships. Thus, the implication of the decision for efficiency and income distribution is most obvious and most salient in this last treatment. This treatment is also the closest to the dictator game designs used in the literature on social preferences (see for a review Fehr and Schmidt, 1999).

7 Appendix 2. Dilution

In this appendix we illustrate how altruism can become diluted in an anonymous partner-selection setting.[18] The trolley experiments (e.g., Greene 2012, Mikhael 2011) suggest that people feel less guilt when the consequences of their action involve mechanical devices, random events, or choices made by others. This suggests that, say, pushing an anonymous person to their death generates more guilt than pushing a button that randomly selects one of M anonymous persons to be pushed to their death. A similar contrast characterizes the difference between treatments T2 and T3: selecting in T2 a partner that leaves other experimental subjects a lower payoff may generate more guilt than indicating in T3 a preference for a partner type that, de facto, removes one partnership from the choice set of others and has similar consequences on randomly selected individuals.

To illustrate how this idea can be formalized, we construct preferences in which individuals value the consequence of their action on others differently depending on whether they affect, with certainty, one person or one of M randomly selected individuals. Let $W_2(h)$ denote the utility gain from choosing a 'high' partner h instead of a 'low' partner l. Consider treatment T2 and let this choice be efficient, so that for a subject with sufficiently altruistic preferences, we have:

$$W_2(h) \equiv \pi_i^h - \pi_i^l + \beta \sum_{j \neq i, j \in N_i} (\pi_j^{h_i} - \pi_j^{l_i}) > 0$$

where $\beta \leq 1$ is a parameter capturing the strength of altruism, N_i is the set of four subjects that includes i, and $\pi_j^{h_i} - \pi_j^{l_i}$ is the effect that player i has on the payoff of player j in N_i when choosing h.

In treatment T3, the effect on the payoff of others is essentially the same as in treatment T2 but i does not know which exact players will be affected. The expected efficiency gain from choosing a

[18] More sophisticated models can be written – e.g., models in which subjects have preferences on whether they interfere or not with other subjects' choices – but they would take us too far from the object of this paper, which is primarily empirical.

partner h can now be written:

$$W_3(h) \equiv \pi_i^h - \pi_i^l + \beta \sum_{j \neq i, j=1}^{24} E[\pi_j^{h_i} - \pi_j^{l_i}]$$

Realized payoffs are as in treatment T2, but the identity of the three individuals affected by i's decision has not yet been determined. It is this difference that opens the door to a possible dilution effect as follows.

Let us first consider i's randomly assigned partner and, without loss of generality, assume that this person belongs to category M. The effect of i's choice on this person's payoff is $\pi_j^{h_i} - \pi_j^{l_i}$; other possible high M partners are unaffected by i's decision so that for them the effect is 0. The total effect of i's choice on the expected payoff of high M subjects is thus:

$$\sum_{j=1}^{6} E[\pi_j^{h_i} - \pi_j^{l_i}] = \sum_{j=1}^{6} \frac{1}{6}(\pi_j^{h_i} - \pi_j^{l_i}) = \pi_j^{h_i} - \pi_j^{l_i} \tag{7}$$

Similar calculations can be done for subjects in the other two category $\times$ type groups. We get:

$$W_2(h) = W_3(h)$$

which predicts that, in the absence of dilution, individuals should make identical decisions under treatments 2 and 3.

Let us now introduce a dilution parameter $\alpha \geq 1$ and rewrite (7) as:

$$\sum_{j=1}^{6} E^{\alpha}[\pi_j^{h_i} - \pi_j^{l_i}] = \sum_{j=1}^{6} \left(\frac{1}{6}\right)^{\alpha} (\pi_j^{h_i} - \pi_j^{l_i}) \leq \pi_j^{h_i} - \pi_j^{l_i} \tag{8}$$

with strict inequality if $\alpha > 1$. The effect on the other two category $\times$ type groups can be handled in the same way. Equation (8) is equivalent to positing that individuals underweight the probability that they affect other players, and is formally similar to probability weighting in prospect theory. With α large enough, $\sum_{j=1}^{6} \left(\frac{1}{6}\right)^{\alpha} (\pi_j^{h_i} - \pi_j^{l_i})$ tends to 0 and we have:

$$W_3(h) = \pi_i^h - \pi_i^l$$

which corresponds to selfish preferences: with enough dilution, people no longer take into account any efficiency cost they impose on others, and pursue their own material welfare only.

8 Appendix 3. Accommodation vs dilution

At first glance it appears that treatments 2 and 3 are fundamentally different in the sense that, in treatment 3, how my choice affect others depends on the actions of others while in treatment 2 it does not. This turns out to be largely an illusion, but showing this rigorously is rather intricate. Although nothing of importance hinges on this issue for our results, we discuss it here in some detail, lest it become a distraction for the reader.

We first show that, if players have the same preferences, treatment 2 and 3 are identical in the sense that my choice affects three other players – yet to be determined, in treatment 3. We then allow for the possibility that players have different preferences, and we introduce the concept of accommodation. Finally we show that, in expectation, there is, if anything, more accommodation in treatment 2 as in treatment 3.

8.1 Developing some intuition

Since all subjects in treatment 3 play the same scenario, players in the same category (A or M) and type (high or low) face the same payoff structure. It follows that, if they all have the same preferences, they should make the same choice. If they make the same choices, the impact on others' payoffs in treatment 3 is the same as in treatment 2: if I take an attractive partner for myself, this attractive partner is not available to someone else and, at the margin, this forces two other players into a less profitable pairing. The only difference is that, in treatment 3, the identity of these players is determined by chance *after* I have made my choice while in treatment 2 the identity of these players is determined by chance *before* I have made my choice. Since the identity of other players is never revealed, the two should be equivalent – except for the dilution effect (see Appendix 2).

If we allow players to have different preferences, I could delude myself into thinking that, while I am acting to maximize my own payoff, others are not – in which case my choice may be *accommodated* by the unselfish choice of others. In other words, I could hope to free ride on another player's altruism.[19] Consequently, in treatment 3 subjects may convince themselves that they are only affecting one participant's payoff as long as there exists other players who exactly complement their choice. Because players face the same payoff structure, such beliefs are largely wishful thinking, however: why would I be the only one who is selfish. But accommodation opens the possibility that players make choices based on their (possibly self-serving) beliefs about how others play.

It therefore appears that how people play in treatment 3 depends on expectations about others' actions while in treatment 2 it does not. This is largely an optical illusion, though. When a player acts selfishly in treatment 2, he may similarly convince himself that others in his group would accommodate his choice – so that he has no determinant effect on the payoffs of other players because these players would accommodate his choice. To illustrate, imagine two grooms and two brides, as in treatment 2. Groom 1 selects the handsomest bride for himself, de facto forcing the other groom to pair with the less desirable bride. Yet groom 1 can convince himself that this is precisely the bride that groom 2 would have selected, no matter how unlikely this possibility may be. In other words, in both treatments 2 and 3 subjects can delude themselves that others will make choices that render their own decision unconstraining.

Having presented the argument in an intuitive manner, we now offer a slightly more rigorous treatment of accommodation, while discussing its relationship with the dilution hypothesis. To recall, we define dilution as the property by which people put less weight on affecting two individuals each with probability 1/2 than on affecting one individual with probability 1. We define accommodation as expecting that other players will accommodate my preferences so that I can pretend to myself that I did not deprive anyone and I can therefore ignore the consequences of my choices on others.

[19] Or, alternatively, I could hope that my altruism is accommodated by someone else's selfishness.

8.2 Accommodation

To formalize the above intuitive argument, we need to distinguish between what i's preferences are with respect to others' payoff, and the realization that other individuals may choose something else than what i would choose in their place, that is, the possibility that others players have different preferences. To demonstrate this, without loss of generality we examine what is i's gain from playing high instead of low:

$$\begin{aligned} W_3^i(h) &\equiv \pi_i^{h_i} - \pi_i^{l_i} + \beta \sum_{j \neq i, j=1}^{N} E[\pi_j^{h_i} - \pi_j^{l_i}] \\ &= M_i^i + \beta \sum_{j \neq i, j=1}^{N} E[M_j^i] \end{aligned}$$

where $\pi_j^{h_i}$ is the payoff to j if i chooses to play high (i.e., to select a high type partner), $\pi_j^{l_i}$ is j's payoff if i plays low, N is the number of players in the game and $M_j^i \equiv \pi_j^{h_i} - \pi_j^{l_i}$.

The whole issue is how $E[M_j^i]$ is calculated. To illustrate how this matters, we now assume that the rules of treatment 3 apply but we vary N. We focus on i, and assume for simplicity that his choice is selected first.

We start by considering the case where $N = 4$ (as in treatment 2). There are three other players apart from i. Let m be the partner selected by i and let k be the player of the same category (A or M) as i but of the opposite type. In treatment 3, the choices of i and k are first collated. If i chooses high and k chooses low, then both their choices are accommodated. Hence player i only directly affects the payoff of one person, i.e., the high player in the other category. In contrast, if i and k choose high, their selections are in conflict and only one of them can be implemented. Imagine that i's choice is selected. In this case, his choice determines the payoffs of all four players. This examples shows that i's choice only affects three other players if k does not accommodate his choice.

Let us now assume that k accommodates i's choice (i.e., chooses low) with probability p. We have:

$$W_3^i(h) = M_i^i + \beta M_m^i + p\beta 0 + (1-p)\beta \sum_{j \neq i, m, j=1}^{N} E[M_j^i]$$

We immediately see that treatment 3 creates an accommodation effect: when player k accommodates i's choice, i can imagine that his choice has no effect on k and the partner that player k chooses. In other words, whatever effect i's choice has on players k and his partner can be 'blamed' on k's choice, not on i's. As this example illustrate, this accommodation effect is present in treatment 2 as well, depending on i's thought process. The key idea is that i can blame someone else for some of the choices made, and this reduces the sense of responsibility i feels for the effect of his choice on others.

This reasoning can be generalized to more players. For instance, consider the case in which $N = 8$ and continue to assume that i plays high and that other k players choose low with probability p. There are three possible choice vectors for the three k players: $\{high, high, high\}$ which happens with probability $(1-p)^3$, $\{high, high, low\}$ which happens with probability $p(1-p)^2$, $\{high, low, low\}$ which happens with probability $p^2(1-p)$, and $\{low, low, low\}$ which happens with probability p^3. We see that if the other three players exactly accommodate i's choice (i.e., the $\{high, low, low\}$ case), i can imagine that his choice has no effect at the margin on k players (and hence on anyone except m) since k players got their choice. In other cases, i's choice has an effect at the margin but only with a certain probability.

8.3 A simple example

This probability can be calculated for a simple example. Imagine $2N$ players asked to choose between two possible prizes, chocolate or marmite, each available in quantity N. Each player expresses a choice for one of them. Let M be the number of players who choose marmite. If there is excess demand for marmite, i.e., if $M > N$, the prizes are assigned at random (as is done in treatment 3) such that the probability of getting marmite is N/M for someone who chooses marmite. Similarly if $M < N$, chocolates are assigned at random. If $M = N$ everyone gets their choice.

Now consider i's choice and imagine that i derives disutility from preventing another player from getting his choice. For simplicity let this disutility be D. Similarly let $-D$ be the utility player i derives from helping another player get his preferred choice. Without loss of generality, imagine that i prefers marmite and let i's payoff in this case be π. Let other players' choices be M_i. If $M_i = N - 1$ then i can get marmite without preventing anyone else from getting their choice. In this case, i's payoff is π. In contrast, if $M_i > N - 1$, by choosing marmite, with probability 1 player i deprives *one* other player from getting their preferred choice. Hence i's payoff is $\pi - D$. Similarly, if $M_i < N$, that is, if most other players do not like marmite, by choosing marmite i makes *one* other player happy with probability 1. Hence i's payoff is $\pi + D$.

Now let p be the probability that other players like marmite. We have:

$$\Pr(M_i = N - 1) = \begin{pmatrix} 2N - 1 \\ N - 1 \end{pmatrix} p^{N-1}(1-p)^N = \frac{(2N-1)!}{N!(N-1)!}(1-p)[p(1-p)]^{N-1}$$

since there are $2N - 1$ players other than i. If $N = 1$ (i.e., there are two players) the above boils down to:

$$\Pr(M_i = N - 1) = \Pr(M_i = 0) = 1 - p$$

With two players only, there is a non-negligible probability that i's choice is accommodated by the other player. In contrast, if N is large, then $\frac{M_i}{2N-1}$ tends to p. In this case, if $p > 0.5$ player i knows that, with a very high probability, he is taking marmite away from people who like it while if $p < 0.5$ he knows he is leaving chocolate for people who like it. In other words, if i has reasons to believe that $p \neq 0.5$ then the *larger* the group is, the *more* likely i is to affect someone else's choice. It follows that if i believes that everybody likes marmite, or that everyone has the same taste as he does, then i's expected utility is $\pi - D$ while if he believes others dislike marmite, his expected utility is $\pi + D$.

If the game's payoff is not a good but money, then player i would normally expect other players to prefer the choice that yields the highest monetary payoff for them. In our example, if marmite is a higher monetary payoff, $p > 0.5$ and player i expects to take away from others' payoff. It follows from the above reasoning that, in a large group where one choice is better for most other players,

i expects to affect other players' payoff such that:

$$\begin{aligned} W_3(h|N &= \infty) \equiv \pi_i^h - \pi_i^l + \beta \sum_{j\neq i, j=1}^{N} E[\pi_j^{h_i} - \pi_j^{l_i}] \\ &\approx \pi_i^h - \pi_i^l + \beta \sum_{j\neq i, j=1}^{N} \frac{4}{N}(\pi_j^{h_i} - \pi_j^{l_i}) = W_2(h) \end{aligned}$$

In contrast, in a small group, say a group of 2, i could imagine that player k will accommodate his choice. This occurs with probability $1-p$. Hence i's utility then is:

$$W_3(h|N=2) \equiv \pi_i^h - \pi_i^l + \beta(1-p)0 + \beta p \sum_{j\neq i, j=1}^{N} E[\pi_j^{h_i} - \pi_j^{l_i}]$$

which is a *less* other-regarding preference, i.e., it is equivalent to a smaller $\beta' = \beta p$. This means that the accommodation effect works in the opposite direction from the dilution effect.

It is possible to show that *for any* p, the probability that other players accommodate i's selfish choice *falls* with N. Run the stata do file below for an illustration:

```
clear all
set obs 101
gen p=(_n-1)/101
forvalues i=1(1)6 {
    gen b`i'=binomialp(2*`i'-1,`i'-1,p)
}
sum b*
scatter b1 b2 b3 b4 b5 b6 p
```

This shows that, if anything, accommodation is *more* likely in treatment 2 than in treatment 3, so accommodation *cannot* be an explanation for dilution.

Conclusion

1. If others' payoff are monetary, then the utility of helping or hindering others' choices can be approximated by the effect on their payoff times a welfare weight β. This simplifies the notation to $W_3(h|N)$.

2. If there is a large group, then the choice of i is nearly never indifferent (i.e., it is nearly never accommodated) *unless* $p = 0.5$ (or more precisely some unique value $\frac{N-1}{2N-1}$ close to 0.5).

3. For groups of intermediate size (e.g., $N = 6$ as in treatment 3) the probability of accommodation is highest for $p = \frac{N-1}{2N-1}$.

4. For groups of smallest size ($N = 1$ as in treatment 2) the probability of accommodation is $1 - p$ and thus is 1 for $p = 0$ and 0 for $p = 1$.

5. The probability of accommodation falls with group size: the larger the group, the lower the probability of accommodation is for any p.

6. If players form rational expectations about others' play, they will expect them to seek to improve their own payoff. Hence $p >> 0.5$ if marmite/high partner is a better choice for the majority of players in i's category and type (who by experimental design share i's payoff structure). In this case, the probability of accommodation is low in all games.

The bottom line is that accommodation *cannot* explain dilution since it operates in the opposite direction.

Table 2 – Scenarios, parameters and choice distribution

	Positive sorting				Negative sorting				Postive sorting efficient	Sharing rule	Share choosing Positive sorting		
Scenario	Bottom 50%		Top 50%		Bottom 50%		Top 50%		Yes / No	Share Low in NS[1]	Bottom 50%	Bottom 50%	Top 50%
	A	M	A	M	A	M	A	M			A	M	A/M
1	10	2	18	18	3	3	15	15	Yes	1/6	0.90	0.09	0.87
2	12	4	16	16	3	3	15	15	Yes	1/6	0.92	0.73	0.80
3	8	8	16	16	9	9	9	9	Yes	1/2	0.00	0.18	0.90
4	10	2	12	12	2.5	2.5	12.5	12.5	Yes	1/6	0.82	0.00	0.41
5	2	2	16	16	5	5	10	10	Yes	1/3	0.00	0.00	0.92
6	6	6	12	12	7.5	7.5	7.5	7.5	Yes	1/2	0.17	0.18	0.96
7	2	2	16	16	9	9	9	9	Equal	1/2	0.09	0.00	0.95
8	7.5	7.5	7.5	7.5	3	3	15	15	No	1/6	0.89	0.91	0.05
9	4	4	11	11	6	6	12	12	No	**1/3**	0.09	0.05	0.17
10	8	2	10	10	9	9	9	9	No	1/2	0.30	0.00	0.73
11	2.5	2.5	12.5	12.5	9	9	9	9	No	1/2	0.03	0.06	0.86
12	8	2	10	10	3	3	15	15	No	1/6	0.90	0.04	0.13
13	2.5	2.5	12.5	12.5	3	3	15	15	No	**1/6**	0.27	0.27	0.09
14	3	3	15	15	8	8	16	16	No	**1/3**	0.09	0.09	0.28
15	8	2	13	13	12	12	12	12	No	1/2	0.92	0.00	0.67
16	9	9	9	9	4	4	20	20	No	1/6	0.08	0.09	0.24
17	3	3	15	15	12	12	12	12	No	1/2	0.00	0.00	0.82

[1] The share is underlined if it is a sharing rule that would guarantee an efficient allocation with selfish preferences.

Table 3. Proportion of choices that maximize the payoff of self or others

	% of choices that maximize own payoff				% of choices that maximize others' payoff			
	T3	T2	T1	χ^2 test	T3	T2	T1	χ^2 test
High payoff subjects	80.8%	84.8%	80.8%	1.93	52.7%	60.0%	68.9%	**13.3**
# of observations	276	282	234	*0.377*	264	270	222	*0.001*
Low payoff A subjects	88.7%	90.9%	85.0%	2.16	43.6%	50.0%	57.5%	**4.87**
# of observations	133	132	120	*0.340*	133	132	120	*0.088*
Low payoff M subjects	90.2%	94.9%	88.9%	3.46	40.2%	42.8%	46.0%	0.91
# of observations	132	138	126	*0.177*	132	138	126	*0.634*
	% of choices that maximize others' payoff: A. when this does not reduce own payoff				**% of choices that maximize others' payoff: B. when this reduces own payoff**			
	T3	T2	T1	χ^2 test	T3	T2	T1	χ^2 test
High types	78.3%	89.5%	94.4%	**17.09**	17.9%	21.4%	36.1%	**10.31**
# of observations	152	153	125	*0.000*	112	117	97	*0.006*
Low A types	83.1%	90.9%	93.2%	3.63	5.9%	9.1%	23.0%	**9.63**
# of observations	65	66	59	*0.163*	68	66	61	*0.008*
Low M types	87.0%	94.8%	92.3%	2.24	7.7%	5.0%	13.5%	3.66
# of observations	54	58	52	*0.326*	78	80	74	*0.160*

T3, T2 and T1 stand for treatment 3, treatment 2, and treatment 1, respectively. The two lower panels break down the upper right panel into cases where the altruistic choice does or does not reduce the subject's own material payoff. The χ2 test columns report the test statistic for equality of means across treatments and, below it in italics, the corresponding p-value of the test. Test statistics significant at the 10% or better appear in bold.

Table 4. Marginal effect of a payoff difference on the probability of choosing high

	T3		T2		T1	
	dy/dx	t-stat.	dy/dx	t-stat.	dy/dx	t-stat.
A. Marginal effect of own payoff difference:						
High type	6.4%	**5.97**	6.6%	**8.41**	6.1%	**7.90**
Low A type	6.6%	**16.64**	6.6%	**11.23**	5.1%	**8.81**
Low M type	5.1%	**4.76**	4.8%	**3.30**	5.4%	**3.26**
B. Marginal effect of others' combined payoff difference:						
High type	0.0%	0.01	1.0%	**2.35**	1.5%	**3.31**
Low A type	-1.5%	**-3.76**	-0.4%	-0.91	0.4%	0.37
Low M type	-0.9%	**-4.84**	-0.7%	-1.36	-0.1%	-0.14

Marginal effects from regression (1).

Table 5. Regression of efficiency on treatment dummies

	All subjects		High payoff		Middle payoff		Low payoff	
	Coef.	**t-stat**	**Coef.**	**t-stat**	**Coef.**	**t-stat**	**Coef.**	**t-stat**
Treatment T1	9.0%	3.64	10.9%	3.22	10.2%	1.84	4.2%	0.93
Treatment T2	6.2%	2.79	5.8%	1.71	10.5%	2.24	2.9%	0.84
Intercept = T3	60.8%	40.03	60.9%	25.99	59.4%	20.97	62.2%	23.72
Nber. Observations	1573		792		385		396	

Dependent variable = percentage of maximum achievable aggregate payoff.

Estimator is a LPM. Standard errors clustered at the participant level.

Table 6. Assignment to archetypes, assuming no mistakes

	All subjects			High types			Low A type			Low M type		
Archetype:	**T3**	**T2**	**T1**	**T3**	**T2**	**T1**	**T3**	**T2**	**T1**	**T3**	**T2**	**T1**
Selfish	57%	64%	45%	54%	53%	38%	50%	73%	55%	68%	78%	48%
Efficient	0%	2%	4%	0%	4%	5%	0%	0%	5%	0%	0%	0%
Equity only	10%	7%	9%	2%	0%	0%	23%	14%	15%	14%	13%	19%
Spiteful	17%	12%	11%	4%	0%	0%	50%	41%	35%	9%	9%	10%
Maximin	22%	23%	21%	2%	0%	8%	18%	14%	20%	68%	78%	48%
Homophily	2%	1%	1%	2%	2%	3%	0%	0%	0%	5%	0%	0%
Multiple archetypes:												
Fits more than one	27%	29%	21%	2%	0%	0%	36%	41%	35%	68%	78%	48%
Fits none	31%	30%	43%	37%	40%	46%	32%	27%	35%	18%	13%	43%
# observations	540	552	480	276	282	234	132	132	120	132	138	126

Table 7. MLE Assignment of subjects to archetypes

	Treatment T3				Treatment T2				Treatment T1			
Pr(Archetype)	**All**	**High**	**Low A**	**Low M**	**All**	**High**	**Low A**	**Low M**	**All**	**High**	**Low A**	**Low M**
Selfish	73.9%	80.9%	57.2%	76.0%	79.6%	82.4%	78.6%	74.7%	59.3%	55.8%	51.5%	73.4%
Efficient	1.1%	0.0%	4.6%	0.0%	4.3%	6.2%	0.8%	3.8%	5.9%	6.7%	7.8%	2.4%
Equity only	1.3%	0.4%	4.3%	0.3%	1.8%	0.0%	6.8%	0.7%	0.0%	0.0%	0.0%	0.0%
Spiteful	15.8%	11.8%	27.1%	12.9%	2.3%	0.3%	1.0%	7.6%	9.1%	12.5%	3.6%	8.1%
Maximin	5.0%	3.5%	6.8%	6.3%	10.9%	8.9%	12.8%	13.1%	14.1%	1.2%	37.1%	16.1%
Homophily	2.9%	3.4%	0.0%	4.5%	1.1%	2.2%	0.0%	0.0%	11.6%	23.8%	0.0%	0.0%
Predicted archetype												
Selfish	81.1%	89.1%	63.6%	81.8%	83.7%	83.0%	86.4%	82.6%	73.8%	64.1%	70.0%	95.2%
Efficient	1.1%	0.0%	4.5%	0.0%	4.3%	6.4%	0.0%	4.3%	5.0%	7.7%	5.0%	0.0%
Equity only	1.1%	0.0%	4.5%	0.0%	1.1%	0.0%	4.5%	0.0%	0.0%	0.0%	0.0%	0.0%
Spiteful	11.1%	4.3%	22.7%	13.6%	2.2%	0.0%	0.0%	8.7%	3.8%	5.1%	0.0%	4.8%
Maximin	3.3%	4.3%	4.5%	0.0%	7.6%	8.5%	9.1%	4.3%	6.3%	0.0%	25.0%	0.0%
Homophily	2.2%	2.2%	0.0%	4.5%	1.1%	2.1%	0.0%	0.0%	11.3%	23.1%	0.0%	0.0%
# observations	540	276	132	132	552	282	132	138	480	234	120	126

Posterior probability of archetype based on mixed MLE model (see text for details). Parameters are estimated separately for each treatment.

Predicted archetype =1 if Pr(archetype)>0.5 and 0 otherwise.