Journal of Public Economics 117 (2014) 104–114

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Social preferences are stable over long periods of time

Fredrik Carlsson [a,*], Olof Johansson-Stenman [a,1], Pham Khanh Nam [a,b,2]

[a] Department of Economics, University of Gothenburg, Box 640, 405 30 Gothenburg, Sweden
[b] Ho Chi Minh City, University of Economics, Viet Nam

ARTICLE INFO

Article history:
Received 20 June 2013
Received in revised form 15 April 2014
Accepted 7 May 2014
Available online 10 June 2014

JEL classification:
C93
H41

Keywords:
Natural field experiment
Preference stability
Social preferences
Voluntary contributions

ABSTRACT

We measure people's pro-social behavior, in terms of voluntary money and labor contributions to an archetypical public good, a bridge, and in terms of voluntary money contributions in a public good game, using the same non-student sample in rural Vietnam at four different points in time from 2005 to 2011. Two of the observed events are actual voluntary contributions (one in terms of money and one in terms of labor), one is from a natural field experiment, and one is from an artefactual field experiment. Despite large contextual variations, we find a strong positive and statistically significant correlation between voluntary contributions, whether correcting for other covariates or not. This suggests that pro-social preferences are fairly stable over long periods of time and contexts.
© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

1. Introduction

The present paper investigates the stability of social preferences by utilizing unique data on people's voluntary contributions to an archetypical public good, a bridge, and contributions in a public good game. The analysis is based on two observations of voluntary contributions, one natural field experiment on contributions to a real bridge in rural Vietnam, and one artefactual field experiment on contributions in a public good experiment, conducted from 2005 to 2011 and using the same sample consisting of all (about 200) households in a village in rural Vietnam. Thus, we obtain repeated information on people's pro-social preferences over a long period of time.

An overwhelming amount of psychological and behavioral economics research shows that the *Homo Economicus* characterization of human behavior, in terms of complete selfishness in a narrow material sense, is often importantly wrong; human behavior is in part pro-social. At the same time, a large heterogeneity in pro-social behavior is typically found. Several studies have consequently attempted to categorize people, based on their experimentally observed behavior, in terms of different types of social preferences, e.g., as free riders, conditional cooperators, and unconditional cooperators (Fischbacher et al., 2001); as selfish versus inequity-averse individuals (Fehr and Schmidt, 1999); and as non-sharers, reluctant sharers, and willing sharers (Lazear et al., 2012). Yet, from these studies one cannot conclude that people are inherently of different types. An alternative explanation is that people simply act differently at different points in time, and that people's degrees of cooperativeness, or non-selfishness, are approximately constant *on average*. Indeed, that people's pro-social actions vary over time is obvious since most of us sometimes contribute to a certain charity and sometimes not. Yet, how much of the observed heterogeneity in social preferences that can be explained by within-people variations is not clear, nor is it clear whether it is significantly more likely that an individual who acted cooperatively at one point in time is more likely to act cooperatively in a similar task several years later. Moreover, even if people are of different types with respect to pro-social preferences, it is an important research issue to find out whether these types are stable over longer periods of time.

The present paper is, to our knowledge, the first in economics to systematically investigate whether pro-social preferences, manifested in terms of cooperative behavior, are fairly stable over several years. In contrast, the extent to which preferences, and in particular social preferences, are stable over a short period of time, and also across decision environments, has been studied in a number of papers with different methods.

Some studies have looked at the differences in pro-social behavior between similar experiments conducted at different points in time. For example, Brosig et al. (2007) conducted dictator and public good

* Corresponding author. Tel.: +46 31 786 41 74; fax: +46 31 786 10 43.
E-mail addresses: fredrik.carlsson@economics.gu.se (F. Carlsson), olof.johansson@economics.gu.se (O. Johansson-Stenman), pham.khanh.nam@economics.gu.se (P.K. Nam).
[1] Tel.: +46 31 786 25 38; fax: +46 31 773 10 43.
[2] Tel.: +46 31 786 47 59; fax: +46 31 786 10 43.

http://dx.doi.org/10.1016/j.jpubeco.2014.05.009
0047-2727/© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

games with the same subjects at several points in time over the course of one week. Other-regarding behavior was found to decrease over time, and in the final experiments the subjects' behavior was close to that predicted by conventional economic theory. Subjects who behaved selfishly were found to be the only ones who behaved stably over time. This pattern is similar to the one typically obtained with repeated public good games.[3] Volk et al. (2012) had subjects participate in a series of three identically designed public good experiments over a period of five months. At the aggregate level, cooperation preferences were stable. However, at the individual level there was considerable variation. Classifying subjects into three categories – conditional cooperators, free riders, and others – it was found that half of the subjects remained in the same category in all three experiments.

Other studies have looked at differences in pro-social behavior between different experiments (with the same subjects). Blanco et al. (2011) ran four different experimental games: dictator, ultimatum, sequential-move prisoners' dilemma, and public good games, and tested whether the Fehr and Schmidt (1999) inequity aversion model can explain the results. They found that the model could explain the results reasonably well at the aggregate level, but that it performs considerably less well at the individual level. De Oliveira et al. (2012) found that preferences for contributing to public goods are positively related across different experimental decision contexts, and also positively related to self-reported donations and volunteering outside the laboratory.

Yet other studies have compared contributions in the lab and in the field. Benz and Meier (2008) conducted a dictator game with two social funds as external recipients and found a positive, albeit relatively weak, correlation between subjects' behavior in a lab experiment and actual charitable giving. Laury and Taylor (2008) found mixed evidence regarding the correlation between non-selfish behavior in laboratory experiments and contribution to a charitable organization. While they found that some measures of altruistic behavior in the lab could be predictive of contributions to the charity, the relationships were generally weak, and some measures of altruism were even negatively correlated with contribution to the charity. Based on a trust game in Peru, Karlan (2005) found that subjects identified as trustworthy, i.e., receivers who returned a relatively large share of what they received from the senders, tend to repay their microcredit loans to a larger extent than those who were not identified as trustworthy in the experiment. No significant correlation between those identified as trusting, i.e., senders who sent a relatively large share to the receivers, and repayment of the loans was obtained. Fehr and Leibbrandt (2011) found that fishermen in Brazil who are more cooperative and patient in lab experiments are also less likely to exploit the common pool resource in the sense that they use shrimp traps with bigger holes (where small shrimp can escape) and fishnets with larger mesh sizes (where only bigger fish are caught). Algan et al. (2013) combine an online experiment with field contribution data for 850 Wikipedia contributors. They found that actual contributions to Wikipedia are strongly related to levels of reciprocity as revealed by both a conditional public good game and trust game. However, they found only weak links with altruism as revealed by dictator game contributions.

Cesarini et al. (2009) used a different approach based on twin studies combined with modified dictator experiments to determine the extent to which giving is heritable. Their best point estimate suggests that genes explain about 20% of the variation in behavior among subjects and hence that social preferences, as manifested in giving behavior in dictator experiments, are in part explained genetically. Yet, this is not necessarily a good measure of the degree to which social preferences are constant over time. First, a certain genetic set-up may in principle induce variation in behavior over time. Second, there are many environmental factors that may work in the direction of stabilizing social preferences, e.g., the development of close relations and social norms.[4]

In summary, most of the studies discussed above point in the direction that social preferences are partly stable over different domains and over time, although the extent of the stability varies between studies to a rather large extent. The main contribution of the present paper is that it investigates the stability of social preferences over long periods of time, where also the contexts differ; we observe behavior in four different periods: in 2005, 2009, 2010, and 2011. As a result of the relatively long time between our events, we argue that we can also more or less rule out the potentially confounding effects of compensatory behavior due to moral licensing and moral cleansing.[5]

The rest of the paper is organized as follows. Section 2 describes the four observations of voluntary contribution and the design of the two experiments, and provides the corresponding background statistics. Section 3 presents the results. We find strong positive and statistically significant correlations between voluntary contributions in these events, whether correcting for other covariates or not, suggesting that pro-social preferences seem to be quite stable over long periods of time.[6] Section 4 concludes the paper.

2. The four observed events of pro-social behavior

We use observations of subjects' pro-social behavior in four related events in 2005, 2009, 2010, and 2011, i.e., implying rather long time periods between events and a total study period of about six years. Two of the events (in 2005 and 2010) are naturally occurring ones where we simply observed the behavior, and the other two (in 2009 and 2011) were designed by the authors. Two of the events (in 2005 and 2009) concern monetary contributions to a local public good in terms of the construction of a crucial bridge in the middle of the village, one of the events (in 2010) concerns labor contributions to construction of the same bridge, while the last experiment (in 2011) was a public good experiment not related to the bridge at all. The first three observations, i.e., in 2005, 2009, and 2010, deal with a real event related to the construction of a bridge. The event in 2009 can moreover be classified as a natural field experiment, whereas the last experiment is an artefactual field experiment, i.e., a lab experiment conducted in the field with a non-student sample.

All four events focus on voluntary contribution mechanisms, and although there are a number of contextual differences, in each instance we observe the behavior of the same (approximately 200) subjects, representing all households in the village. The events were undertaken in the Giong Trom hamlet,[7] in the Mekong river delta of Vietnam, where

[3] See, e.g., Isaac et al. (1984), Andreoni (1995), and Fehr and Gächter (2000). Different explanations have been proposed, including initial confusion and learning (e.g., Andreoni, 1988) and some version of conditional cooperation (e.g., Fischbacher et al., 2001; Fischbacher and Gächter, 2010).

[4] There is also a large (although often questioned) literature in psychology arguing that there are five basic dimensions of the human personality: openness, conscientiousness, extraversion, agreeableness, and neuroticism. The empirical analysis in this literature is most often based on surveys with verbal questions trying to capture these dimensions, and it is often argued that individual measures along these dimensions are rather stable over time. For example, Jang et al. (1996) found, based on a twin sample, that approximately 40% of the variation in agreeableness, which reflects a tendency to be cooperative, trusting, and helpful rather than antagonistic, can be explained genetically.

[5] Moral licensing (e.g., Monin and Miller, 2001; Mazar and Zhong, 2010) suggests that people who have undertaken a praiseworthy act get an implicit license for subsequently conducting a more selfish act, whereas moral cleansing (e.g., Carlsmith and Gross, 1969; Tetlock et al., 2000) refers to the flip side where a morally blameworthy act induces compensatory behavior. Thus both mechanisms work in the direction of reducing the correlation between pro-social actions over time. In the present case, we argue that, unlike most previous tests of the stability of social preferences, the time periods are sufficiently far away such that the effects of moral licensing and cleansing are presumably small.

[6] This does of course not mean that pro-social behavior is independent of the social context. For example, donations to charitable organizations have been shown to depend on the information about what other people donate (see, e.g., Bardsley and Sausgruber, 2005; Alpizar et al., 2008; Frey and Meier, 2004; Shang and Croson, 2009) and on whether the action is observed by others or not (Soetevent, 2005; Hoffman et al., 1996; List et al., 2004).

[7] A hamlet is a small village or part of a village and consists of around 100 to 300 households.

about 200 households live and use the bridge (if/when it is in sufficiently good shape). Most households in the hamlet are engaged in rice cultivating activities.[8] The hamlet suffers from a problem that is common in the Mekong river delta: lack of basic infrastructures such as rural roads, bridges, and irrigation canals. The government only provides larger public goods such as roads between villages, whereas small-scale infrastructures within a hamlet are considered to be the responsibility of the hamlet.

2.1. The bridge and the voluntary contributions

The bridge is important for the village because villagers use it to go to the rice fields, to the market, to school, and to visit friends, given that the bridge is in sufficiently good condition. If they do not use the bridge, they have to choose one of two other routes, each located parallel to and about 1200 m from the bridge's pathway; see the map in Fig. 1. There are people living on both sides of the canal, and there is also important infrastructure on both sides of the canal. For example, the main market is located on the north side of the canal, while the school is located on the south side. For all but a few households, the travel distance to reach the destinations on the other side is considerably longer if they cannot use the bridge.

The first three observations of pro-social behavior, to be described next, concern funding of a bridge for the hamlet.

2.2. The 2005 event

In the first event in 2005, the hamlet council had decided to try to build a bridge, which was to be funded by voluntary contributions. A group of three delegated individuals visited every household in the hamlet to present the plan to build the bridge and to ask for voluntary contributions. Probably in order to persuade villagers and increase contributions, the delegates showed a list of names, contribution amounts, and signatures of those who had already contributed. The hamlet council did not set an upper contribution limit, and the highest contributed amount was 300,000 dong.[9] Since the total contribution was not sufficient to build a concrete bridge, the hamlet council decided to build a wooden bridge. Yet, the bridge became degraded relatively quickly, and in 2009 it looked like in the picture in Fig. 2 and could obviously not be used for tractors and motorbikes.

The full list of contribution amounts from each household was kept by the hamlet council and made available to the authors of this paper.

2.3. The 2009 experiment

In collaboration with an NGO, we conducted a field experiment using a threshold public good game that concerned the funding of a new bridge for the hamlet in 2009. Fifteen experimenters were recruited via advertisements and a careful selection procedure at the University of Economics Ho Chi Minh City. The selected experimenters received extensive training and spent nearly one week practicing the experiment in role-play pairs and for pilot interviews with farmers. Moreover, the selected experimenters also had experience from similar fieldwork in a neighboring rural area. A list of likely questions and answers related to the project was provided and the experimenters were repeatedly told about the importance of using the exact wordings prescribed in the script. For a detailed description of the experiment and the results, see Carlsson et al. (2013).

The main objective of the experiment was to investigate the role of social influence for voluntary contributions to public goods. We devised a threshold public good game, in which each of the 200 households received a 400,000 dong endowment from the NGO and had the option to either keep the money or contribute some or everything to the funding of the bridge. The threshold level was set at 40 million dong, meaning that if all villagers together would contribute at least 40 million dong, the bridge would be built; otherwise it would not. The threshold was explained as follows:

> *The concrete bridge will be established if all families together contribute at least 40 million dong. This means that if the total contribution is equal to or above 40 million dong, the project will use these 40 million dong, add more funding in order to meet the costs of the bridge, and take the responsibility to build the bridge. If the total amount of money collected exceeds 40 million dong, the excess amount will be returned to your family according to the proportion you contributed.If the families are unable to contribute a total of 40 million dong, your contribution will be returned to you, and the concrete bridge will not be built.*

The experiment involved five randomly distributed treatments as follows: 1. Baseline case with no reference contribution level and no default option. 2. High reference contribution level (300,000 dong) and no default option. 3. Low reference contribution level (100,000 dong) and no default option. 4. No reference contribution level and a default option at zero contribution, and 5. No reference contribution level and a default option at full contribution (400,000 dong) of the endowment.

The reference contribution levels were conducted by providing the subjects with information about a typical previous contribution of others. These numbers, in turn, were obtained from an initial treatment where we did not tell the subjects anything about others' contributions. The default options were conducted using a metal card with 9 different contribution levels. A magnetic token was initially put at the 0 dong level or at the 400,000 dong level. The subjects were then asked to move the token to the amount that they wanted to contribute to the public good. In all treatments, the contributions were anonymous to everybody except the solicitors, i.e., the contributions were not revealed to any parties. Since the households contributed enough to reach the threshold, the new bridge was built in early 2010; see the picture in Fig. 3.

2.4. The 2010 event

During the planning of the construction, we had a meeting with the head of the hamlet and representatives from the farmers' association. At the meeting, we were informed that they planned to ask the villagers to contribute labor to connect the road with the new bridge. We took this opportunity to collect another naturally occurring contribution data set. As the construction work required joint efforts in a short period of time, three and a half day were scheduled for the joint work. Two persons from the hamlet council visited the households in the hamlet to invite villagers to contribute labor. Hence, an important difference compared with the previous two events is that instead of being asked for monetary contributions, they were asked for labor contributions.

Since some households were not expected to be able to contribute any labor at all, mainly because the household members were too old, not all households were asked to make contributions. In total 19% of the households were not asked to make any labor contributions.[10] At this time, households were not told anything about what others were contributing, and there was no provision point. We hired an external supervisor who monitored the construction progress and quality, and recorded villagers' labor contributions. Thus, what we observe in this event is the actual amount of labor contributions, and not what they promised when asked to contribute.

[8] There are slightly more than 200 households that administratively belong to the hamlet. However, a small group of households live separated from the rest and have no use of the bridge at all, and were not included in any of the events.

[9] At the time of the experiment, 100,000 dong = 5 USD.

[10] Estimating a binary probit model where the dependent variable is equal to one if they were not asked to contribute, we find that small and poor households and households with an old head or a female head were more likely not to be asked. This is all as expected.

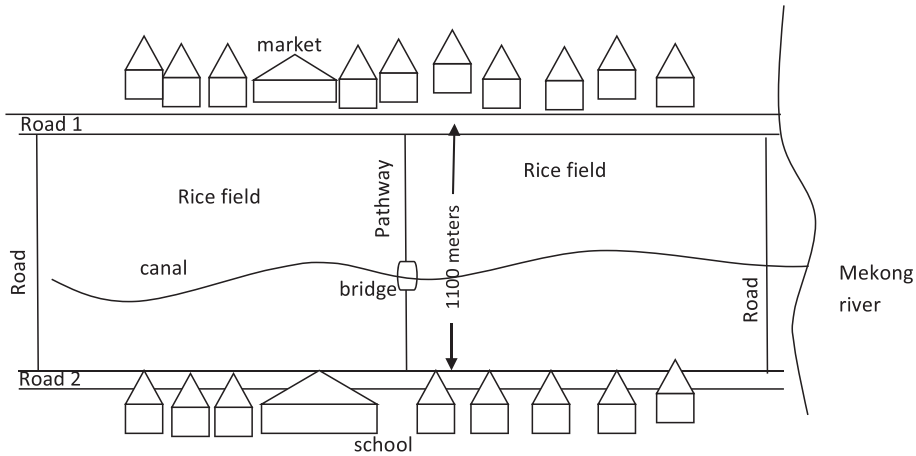

Fig. 1. Map of the studied village.

2.5. The 2011 experiment

While the bridge is presumably useful to all households, the usefulness varies with, e.g., the distance to the bridge and ownership of different vehicles. Although we have information on the use of the bridge, it is possible that we still cannot perfectly correct for it in our analysis. In order to avoid such potentially confounding effects, it makes sense to also include an experiment that by design is not at all related to the bridge.

The experiment conducted in 2011 was therefore not in any way related to the bridge or the 2009 experiment, and no reference to either the bridge or the earlier experiment was made. While the 2009 experiment was funded by an NGO, the 2011 experiment was conducted and funded by the University of Economics Ho Chi Minh City. There were 15 experimenters also in this experiment. They were recruited and trained in a similar manner as for the 2009 experiment, but none of them were the same as in the 2009 experiment. The experiment was conducted in the subjects' homes and the first part of the experimental script read:

> *Hello, my name is […] I'm working for University of Economics Ho Chi Minh City, which is implementing an economic study in this village. I would like to invite you to participate in an economic experiment, which will be conducted now.You will be taking part in an experiment on decision-making and will earn some money by choosing among alternatives. The experiment is designed so that your earnings will depend on your choices, as well as on the choices of others. Your earnings will be paid to you in cash tomorrow afternoon privately to preserve the confidentiality of your earnings. The results from this study will only be used for academic purposes, and no other people in the village will have information about your choices.*

We chose to conduct a public good game, much like a standard linear public good game conducted in laboratory experiments. Yet, in order to fit the setting of the village, and in order to be able to easily compare the contribution behavior in this experiment with the other experiments, the group consisted of all households in the village; thus, the size of the group was approximately 200 subjects. Each of these households received 200,000 dong, which was clearly a substantial amount for them. Just as in a standard laboratory public good experiment, they had to decide how much of the endowment to keep and how much to put into a group account. In order to make the experiment simple to understand, each subject was told that any money that was put into the group account would be doubled by the experimenter, and that the total amount in the group account would be distributed to the group members, and hence to the households. The following sentences were part of the script that the experimenters read to the subjects.

> *This experiment is organized with a group of 200 households including yours. We will give 200,000 dong to each of the 200 households participating in this experiment. Here is the agreement, which has a signature and stamp from the University of Economics Ho Chi Minh City stating that 200,000 dong belongs to your household. Your household and each of the other households have to decide how much of the 200,000 dong to keep and how much to put into a group account.The amount of money you put into a group account will first be doubled by the experimenter, meaning that if you transfer 1 dong into the group account, it will become 2 dong. The total amount will then be divided equally between all households.*

The instructions included several examples, and the subjects were allowed enough time to understand everything that was said. To inform

Fig. 2. The wooden bridge built from the 2005's contributions.

Fig. 3. The concrete bridge built from the 2009's contribution.

Table 1
Characteristics of the four events.

Characteristics	Event 1	Event 2	Event 3	Event 4
Time	July 2005	August 2009	March 2010	September 2011
Contribution mechanism	Voluntary	Voluntary	Voluntary	Voluntary
Anonymity	No	Yes	No	Yes
Framework	Fundraising campaign	Threshold public good game	Voluntary contributions	Public good game
Windfall money	No	Yes	No	Yes
Contribution range	[0, .) thousand dong	[0, 400] thousand dong	[0, 3.5] labor days	[0, 200] thousand dong
Organizer	Local government	Outside NGO	Local government	University
Reference contribution	Yes	Yes in some treatments	No	No

the subjects that their contributions would be kept completely confidential, the experimenters told them before they made their contribution decision: "No one in the commune, not even the officials, will know about your decision. We will keep your contribution information completely secret."

2.6. Summary of the four events and household characteristics

As mentioned, the first three events were related to the bridge in the hamlet. The first event concerned monetary contributions to build a small wooden bridge in 2005; the second concerned monetary contributions to build a new and better concrete bridge in 2009; and the third event concerned voluntary labor contributions to connect the road to the new concrete bridge in 2010. The fourth event was instead a public good experiment that was not related to the use of the bridge. The settings of the four events are summarized in Table 1.

While we designed only two experiments (the 2009 and 2011 experiments), we have data for four different points in time, i.e., 2005, 2009, 2010, and 2011, for essentially the same subjects. Table 2 reports background statistics, as of 2009, for the households. There were a total of 200 households from the village that participated in the events. However, as explained above, not all households were asked to contribute labor in the 2010 event, and in the last experiment four of the households were unable to participate.[11]

The mean monthly household income is around 1.8 million dong per month. This amount corresponds to about 95 USD per month, which is less than one USD per household member and day. Thus, the households in the study are poor, and the average level of education is very low. The average size of land on which a family is currently cultivating rice is also rather small, approximately half a hectare.

3. Results

3.1. Average contributions in the four events

Before looking at the correlations between the contributions, let us briefly look at Fig. 4, which displays the histograms of contribution in each of the events.

As can be seen, the contribution patterns are strikingly different, in particular between the 2005 and 2010 events, on the one hand, and the 2009 and 2011 events on the other. The 2009 and 2011 events have in common that they are based on windfall endowments, which may contribute to the substantially higher contribution levels in these events.[12] Comparing the 2005 and 2009 contributions, which were both in terms of monetary contributions to a new bridge, there are thus strikingly large differences. The mean contribution in 2009 was almost seven times as large as in 2005 (270,000 dong compared with 40,000 dong), and while almost everyone contributed something in 2009, almost half of the households chose to free ride in 2005. While there may be many different explanations for this, two clearly stand out: First, as mentioned, the contributions in 2009 were based on a windfall endowment provided by the NGO, while the contributions in 2005 were paid out of the households' existing wealth. Second, the 2009 experiment involved a matching contribution by the involved NGO. Such matching contributions, or seed money, have been shown to increase voluntary contributions substantially (e.g., List and Lucking-Reiley, 2002; Karlan and List, 2007).

Moving on to the 2010 event, we can observe that even fewer chose to contribute than in 2005, in this case in terms of labor contributions. In 2010, the mean contribution of labor was 0.5 labor days per household, which corresponds to about 40,000 dong based on an average daily labor wage of 80,000 dong. Finally, in the 2011 experiment, there is again a small share of people contributing nothing, where the mean contribution is substantial, 125,000 dong. The mean contribution as a fraction of the maximum contribution is actually quite similar to the 2009 experiment; these events also share the features that they are concerned with voluntary financial contributions to a public good and that the contributions are based on windfall endowments.

The contribution levels in the 2011 experiment are clearly unusually high compared with what is typically found in public good lab experiments, in particular given the small marginal per capita return. There are several potential reasons for this. Contributions in one-shot public good experiments are typically larger than average contributions in multi-round public good experiments (e.g., Fischbacher and Gächter, 2010). Also, there is some evidence that non-student samples might be more cooperative and show higher levels of reciprocation than student samples (e.g., Gächter et al., 2004; Falk et al., 2013). Moreover, the participants in this particular non-student sample may know each other better than most students participating in lab experiments, and there is (not surprisingly) evidence that subjects tend to behave more pro-socially when the social distance to the other subjects is small (e.g., Hoffman et al., 1996). Finally, it is also possible that the fact that the experiment took place in the subjects' homes may have induced stronger psychological pressure to contribute than if it had been undertaken in a common venue, which appears natural given the evidence that subjects' contributions tend to be affected by whether or not the experimenters directly observe the contributions (List et al., 2004; Alpizar et al., 2008).

Yet, we are not primarily interested in the various contribution *levels*, or whether these levels are similar across decisions. Instead we want to find out to what extents decisions are *correlated*, i.e., whether or not those households that contribute more in one event also tend to contribute more in another.

3.2. Raw contribution correlations between the events

As described above, we observe the contributions in each of the events at the household level. As a first step, we therefore analyze the simple pair-wise correlations between the four events. Remember that, for each household, we have three observations of contributions to the bridge and one observation of contribution in a public good experiment. We present correlation coefficients for the whole sample and for the restricted sample of households that had the possibility to

[11] One of the households was attending a funeral, and in the other three households the household head was working outside the village at the time of the experiment.

[12] There are a few tests of the effects of windfall endowments in public good experiments. Cherry et al. (2005) and Clark (2002) found no evidence of a windfall-gain effect on contributions, while Kroll et al. (2007) found significant differences in a public good experiment with heterogeneous endowments.

Table 2
Household characteristics.

Variables	Definitions	Mean	Std. dev.
Household size	Number of household members	3.84	1.61
No labor	=1 if household cannot provide labor for community work	0.19	
Age	Age of household head in year	48.9	13.8
Male	=1 if household head is male	0.63	
Education	Highest level of education attained: 1 = no schooling (5%); 2 = grades 1–5 (54%); 3 = grades 6–9 (31.5%); 4 = grades 10–12 (9%); 5 = vocational school and above (0.5%)	2.46	0.76
Monthly income	Monthly household monetary income in hundred thousand dong	18.13	12.78
Use the bridge everyday[b]	=1if use bridge everyday	0.19	
Use the bridge 1–3 times a week[a]	=1 if use bridge about 1–3 times a week	0.10	
Use the bridge twice a month	=1 if use bridge about 2 times a month	0.17	
Use the bridge once a month	=1 if use bridge about 1 time a month	0.30	
Member of the communist party	=1 if at least one household member is a member of the communist party	0.10	
Member of a local association	=1 if at least one household member is a member of a local association	0.49	
Rice land	Total size of rice land currently being cultivated; in congs (1 cong = 1/10 ha)	4.69	3.13

[a] The options for the question regarding current use of the bridge were: 1 = Every day, 2 = About two to three times a week, 3 = About once a week, 4 = About twice a month, 5 = About once a month or less, 6 = Currently do not use the bridge at all. Since relatively few respondents chose options 2 and 3, we merged them in the descriptive statistics and in the analysis.

[b] We also have a measure of the physical distance from each home to the bridge. However, this measure never contributed significantly to explain contributions to the bridge. There are several potential reasons for this: For example, some farmers live on the same side of the bridge as their rice field, while others live on the opposite side, and the same applies for households with kids related to the position of the school. Hence, we do not include that variable in the analysis.

contribute in 2010. For those that were not asked to make labor contributions, we set the contribution to zero when calculating the correlations for the whole sample. Table 3 presents the pair-wise correlation coefficients.

Despite the large differences in contribution levels between the events, including in the fraction that did not contribute anything, the correlation coefficients between the four events are substantial and in most cases highly significant. For the events related to the bridge, the largest correlation coefficients are found between the 2005 and 2010 events. This may seem surprising since the 2005 event concerns monetary contributions, while the 2010 event concerns labor days. Also, it seems likely that some people have a comparative advantage in labor contributions, implying that there is scope for a degree of specialization in contributions, which should reduce the correlation coefficient. Our favorite explanations for the large correlation coefficient are as follows: First, in both cases the subjects had to pay with their own resources (money and time, respectively), and hence there were no windfall resources obtained for the individual decision. Second, and perhaps even more importantly, both of these events were non-anonymous, and if some people are more sensitive to the peer pressure to contribute, then they should contribute more than others in both instances, implying a positive effect on the correlation coefficient. Indeed, Alpizar et al. (2008) demonstrate in a field setting that anonymity matters for charitable contributions and Andreoni and Rao (2011) found more recently that communication per se seems to dramatically affect altruistic behavior, primarily through increased empathy.

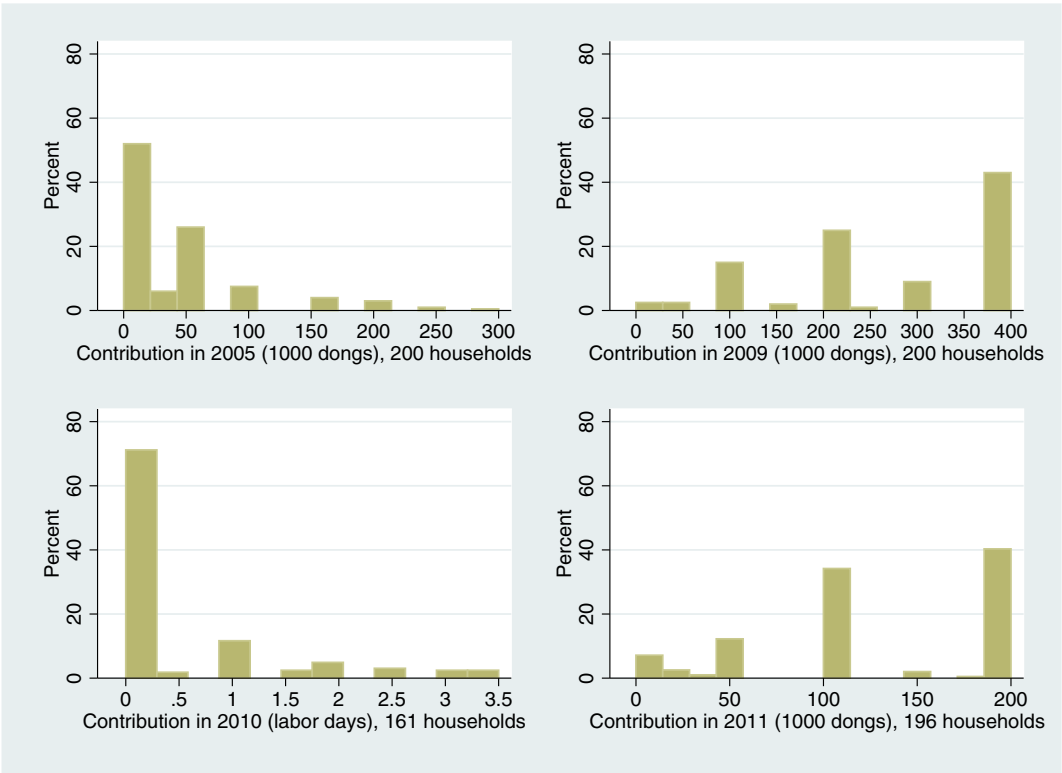

Fig. 4. Histograms of contributions in the four events.

Table 3
Correlation coefficients between contributions for the different events.

	Whole sample (N = 196)			
	Contribution 2005	Contribution 2009	Contribution 2010	Contribution 2011
Contribution 2005	1.00			
Contribution 2009	0.28***	1.00		
Contribution 2010	0.37***	0.22***	1.00	
Contribution 2011	0.20***	0.28***	0.14**	1.00
	Restricted sample (N = 161)			
	Contribution 2005	Contribution 2009	Contribution 2010	Contribution 2011
Contribution 2005	1.00			
Contribution 2009	0.25***	1.00		
Contribution 2010	0.41***	0.26***	1.00	
Contribution 2011	0.22***	0.30***	0.12	1.00

*, **, and *** denote that the coefficient is statistically significant at the 10%, 5%, and 1% levels, respectively.

The correlation coefficients between the experiment in 2011 and the other three events are also substantial, with the exception of the event involving labor contributions in 2010. Here, the coefficient is not statistically significant at conventional levels when based on the restricted sample. There may be several reasons for this, in addition to the fact that only one of the two is related to the bridge: one concerns labor contributions while the other concerns monetary contributions, one is anonymous while the other is not, and one was conducted based on windfall money while the other was not. Yet, it is interesting to see that the correlation coefficients between the contributions in 2005 and 2011 and between the contributions in 2009 and 2011 are large and statistically significant, despite the large differences in set-up. Together, this clearly shows (*i*) that the strong correlations between the events cannot only be due to the fact that they concern contributions to a similar good, i.e., the bridge, and (*ii*) that there is clear support for the idea that social preferences reflect traits that to a large extent are constant over time and domains.

3.3. Simple comparisons of Low, Intermediate and High contributors' behavior over time

Correlation coefficients are not always straightforward to interpret, in particular as in our case where the contribution patterns are very different between the events and where large fractions of the subjects chose the extreme alternatives of contributing zero or the full amount. We therefore also illustrate the stability of social preferences by investigating how subjects who initially contributed small or large amounts in 2005 contributed in the three subsequent events. Based on the histograms above it appears natural to classify those who contributed strictly more than 100,000 dong as High contributors, implying that 9% of the subjects are characterized as High contributors in 2005. Since as many as 47% of the subjects contributed zero in 2005, this constitutes a natural (albeit quite large) group of Low contributors (implying that the remaining 44% are classified as Intermediate contributors).

Table 4
Low, Intermediate and High contributors in the events 2009, 2010, and 2011 for those who were characterized as Low, Intermediate and High contributors in 2005.

	Sample based on Low contributors in 2005[a] (n = 93)	Sample based on Intermediate contributors in 2005[b] (n = 90)	Sample based on High contributors in 2005[c] (n = 17)
Contribution in 2009 (dong)			
Low contributor (≤100,000)	23	17	0
	(25%)	(19%)	(0%)
Intermediate (>100,0000 & ≤400,000)	39	33	2
	(42%)	(37%)	(12%)
High contributor (=400,000)	31	40	15
	(33%)	(44%)	(88%)
Contribution in 2010 (labor days)			
Low contributor (=0)	84	65	4
	(90%)	(72%)	(24%)
Intermediate (>0 & <2)	5	14	7
	(5%)	(16%)	(41%)
High contributor (≥2)	4	11	6
	(4%)	(12%)	(35%)
Contribution in 2011 (dong)			
Low contributor (≤50,000)	28	16	1
	(30%)	(18%)	(6%)
Intermediate (>50,0000 & <400,000)	35	39	2
	(38%)	(43%)	(12%)
High contributor (=400,000)	30	36	14
	(32%)	(39%)	(82%)

[a] Those who contributed 0 dong 2005.
[b] Those who contributed more than 0 but less than or equal to 100,000 dong 2005.
[c] Those who contributed more than 100,000 dong 2005.

The subjects for the three subsequent events were similarly classified as Low, Intermediate or High contributors. For the event in 2009, Low contributors are those contributing less than or equal to 100,000 dong (20% of the subjects) whereas High contributors are those who contributed the maximum amount 400,000 dong (43% of the subjects). For the event in 2010, Low contributors are those who did not provide any labor days (77% of the subjects) and High contributors are those who provided two or more labor days (11% of the subjects). Finally, for the event in 2011, Low contributors are defined as those who contributed less than or equal to 50,000 dong (22% of the subjects) and High contributions are those who provided the maximum contribution 200,000 dong (40% of the subjects).

Table 4 reports the shares of Low, Intermediate and High contributors in the events 2009 to 2011 based on three different samples, those who in 2005 are classified as Low, Intermediate, or High contributors, respectively.

In general, we find strong links between contributing behavior in the 2005 event and the behavior in the subsequent events. Since the overall fractions of Low and High contributors are different from each other in each of the experiments the relevant comparisons to make are between the different samples, i.e., to compare the different shares horizontally.[13] That is, we want to compare whether those who in 2005 are classified as a High contributors are more likely to be classified as a High contributors in the three subsequent events than what those who are classified as a Low contributor in 2005 are. Correspondingly, we want to compare whether those who in 2005 are classified as a Low contributor are more likely to be classified as a Low contributor in the subsequent events than what those who are classified as a High contributor in 2005 are. As can be seen in Table 4, the pattern follows expectation for each of the six cases.

For example, it is striking to compare the non-anonymous voluntary contributions to the bridge in 2005 with the contributions in the anonymous public good game (not related to the bridge) six years later. Of those classified as Low contributors 2005, 32% are classified as High contributors 2011. This can be compared with those classified as High contributors in 2005, where as many as 82% are classified as a High contributor 2011. Similarly, of those classified as Low contributors 2005, 30% are also classified as Low contributors 2011. This can be compared with those classified as High contributors in 2005 where as few as 6% are classified as a Low contributor 2011.

In order to formally test whether these differences are statistically significant, we focus on the two samples of Low and High contributors in 2005. We can then use simple Chi-square tests when comparing the distribution of Low, Intermediate and High contributors for each of the 2009, 2010 and 2011 events. For example, we can test if there is no difference in the distribution of Low, Intermediate and High contributors in the 2009 event for those that were Low contributors in 2005 and those that were High contributors in 2005. In all three cases, i.e. for each of the events in 2009, 2010 and 2011, we can reject the null hypothesis that the distribution is independent of whether a subject was classified as a Low or High contributor in 2005, at the 1% level.[14]

3.4. Econometric analysis

While the strong positive correlation coefficients, as well as contribution patterns more generally, between contributions in the first three events (i.e., those related to the bridge) are interesting per se, one should be hesitant to interpret these coefficients as clear evidence of stability of social preferences. Indeed, there are several possible interpretations of these positive correlations. For example, if the households using the bridge the most are also willing to contribute the most (e.g., for partly selfish reasons), we should obtain a positive correlation between contributions in the first three events even if there are actually no differences between the households in terms of underlying social preferences. Still, the use of the bridge can of course not explain the correlation coefficients between contribution in the public good experiment in 2011 (which was not related to the bridge) and contributions in any of the other events.

We deal with this problem in two ways: First, as mentioned, we run a public good experiment without any reference to the bridge, based on the same sample. Second, we use regression techniques correcting for possible explanatory variables that can be assumed to vary across the households but at the same time are independent of underlying differences in social preferences. The most obvious variable here is the use of the bridge.

More specifically, we use multivariate tobit regressions since we have non-negligible shares of subjects who either contribute the full amount or do not contribute at all; hence, we use truncations at both zero and the full amount for the experiments in 2009 and 2011, and at zero for the events in 2005 and 2010. Using a multivariate model, we estimate the correlation coefficients of the error terms for each event. These error terms are assumed to reflect the part of social preferences that cannot be explained in terms of our explanatory variables used in the regressions. Moreover, simple correlations do not take into account that there were different treatments in the experiment in 2009. In order to deal with these issues, we estimate multivariate models where the four equations are estimated simultaneously, allowing for a correlation between the error terms of each of the equations, and where the dependent variables are censored.

We present three sets of regressions: In the first set, we use no explanatory variables (except for an intercept). In the second set, we use only variables reflecting the use of the bridge in the first three events, since these variables presumably vary across the households and at the same time are independent of underlying differences in social preferences. For the second and fourth event we also include treatment dummy variables and experimentalist dummy variables. Finally, we present a third set of regressions including all relevant explanatory variables. In this last set, we thus face the risk of "over-compensation" in the sense that there may exist variables, such as age and income, that are correlated with true underlying social preferences. For example, suppose that all actual variations in social preferences are determined by gender. If we then correct for gender in the regressions, the result would indicate that there is no stability of social preferences over time even though the actual stability through gender may be large. Yet, as is the case when not including any explanatory variables, it constitutes a natural benchmark case.

We focus mainly on the sample of households that had the possibility to contribute labor in 2010. However, we also report the results based on the full sample, where we have hence set the contribution in labor to zero in 2010 for those who were not asked to contribute. We also use the full dataset for the pairwise correlations that do not include the 2010 event. Yet, as can be observed, the results turned out fairly similar.[15] The estimated correlation coefficients for our three sets of multivariate regressions are presented in Table 5, for each separate event.

Starting with the first three events, we can observe that the pairwise correlation coefficients are consistently positive, substantial, and statistically significant. Consequently, even when controlling for a number of

[13] For example, it does not say that much that among those who are classified as High contributors in 2005, 88% are classified as High contributors in 2009, and no one is classified as Low contributors in 2009, since there are overall more subjects in 2009 that are characterized as High contributors than who are characterized as Low contributors. Similarly, it does not say much that among those who are classified as Low contributors in 2005 there are actually slightly more subjects in 2011 who are classified as High contributors than who are classified as Low contributors, since there are overall more people in 2011 who are characterized as High contributors than who are characterized as Low contributors.

[14] Chi-squared numbers and corresponding probability figures for each of the three cases 2009, 2010 and 2011 are 18.07 (p-value = 0.000), 40.08 (p-value = 0.000) and 15.07 (p-value = 0.001) respectively.

[15] We also estimated a bivariate tobit model where we only included the monetary contributions in 2005 and 2009 based on the full sample of 200 subjects. The results do not differ in any substantial way compared with what we will present in the main text here and are thus not reported, but are available upon request.

Table 5
Estimated pairwise correlation coefficients between the error terms from multivariate tobit regressions (number of draws = 200); dependent variables are contributions in the four events.

	No variables (except intercept)				Only use-the-bridge variables				All variables			
	2005	2009	2010	2011	2005	2009	2010	2011	2005	2009	2010	2011
Treatment dummy variables	No	No	No	No	No	Included	No	No	No	Included	No	No
Experimentalist dummy variables	No	No	No	No	No	Included	No	Included	No	Included	No	Included
Socio-economic variables	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Restricted sample (N = 161)												
2005	1				1				1			
2009	0.36***	1			0.33***	1			0.30***	1		
	(0.08)				(0.09)				(0.09)			
2010	0.49***	0.33***	1		0.47***	0.25**	1		0.41***	0.19*	1	
	(0.08)	(0.09)			(0.08)	(0.10)			(0.09)	(0.11)		
2011	0.28***	0.33***	0.16	1	0.22**	0.37***	0.04	1	0.22**	0.39***	−0.04	1
	(0.09)	(0.08)	(0.10)		(0.10)	(0.09)	(0.12)		(0.10)	(0.10)	(0.13)	
Whole sample (N = 196)												
2005	1				1				1			
2009	0.38***	1			0.33***	1			0.29***	1		
	(0.07)				(0.08)				(0.08)			
2010	0.49***	0.30***	1		0.47***	0.25**	1		0.42***	0.19*	1	
	(0.08)	(0.09)			(0.08)	(0.09)			(0.09)	(0.10)		
2011	0.26***	0.32***	0.19**	1	0.22**	0.30***	0.08	1	0.19**	0.28***	0.01	1
	(0.08)	(0.08)	(0.09)		(0.08)	(0.09)	(0.11)		(0.09)	(0.09)	(0.11)	

*, **, and *** denote that the coefficient is statistically significant at the 10%, 5%, and 1% levels, respectively.

observable differences among households and the treatment effects, there are strong correlations in behavior between the three events. The relative sizes of these coefficients follow expectations in that they are generally largest when we do not correct for any variables, and smallest when we include the full set of variables. Yet, the differences between when we correct for the use-of-the-bridge variables and when we do not are rather small. Again, the highest correlation coefficients are found between the 2005 and 2010 events, probably largely due to the fact that these events were not anonymous, as discussed previously.

Yet, it could still be the case that we did not manage to completely correct for the usefulness of the bridge to different households, and that the remaining part could explain the positive correlations. For this reason, we have the fourth artefactual field public good experiment, which is completely unrelated to the bridge. Here we again find a weak effect between 2010 and 2011 (we speculate about possible reasons above), while the large and highly significant correlations between contributions in 2005 and 2011 and between contributions in 2009 and 2011 largely prevail after including various covariates. Consequently, we can again conclude that it is not the fact that three of the four events concerned the same underlying good, a bridge that explains the significant correlation between the contribution decisions.

Consider next the estimated coefficients for the covariates. Table 6 presents the results for the restricted sample. Few of the household characteristics have a significant impact on the contributions in any of the events and there are not any consistent patterns across the different events (except for the use of the bridge). The contributions in 3 out of the 4 events, including the public good experiment in 2011, are positively related to the amount of rice land owned. Monthly income, in contrast, is only positively correlated with contribution in the public good experiment. One may speculate about the reason for the difference. Perhaps it is more natural to relate to own income when making the contribution decision in a non-framed public goods experiment than when making decisions that are linked with the bridge. For the 2010 event, when people made contributions in terms of labor time, it is not surprising that there is a negative effect of age. It is less expected to find a negative effect of the household head being male. As suggested by a referee, a rough, indicative story can be given as follows: One does not have to use the bridge very much to give out of the windfall, one needs to use it a little more to give labor (free), and one needs to use it a lot to give money if poor.

3.5. Are the obtained correlation coefficients large?

Despite the differences in context, which we know are likely to affect the contribution levels, we find substantial and in most cases statistically significant correlations among the four contribution decisions. Are these correlation coefficients large? We argue that they are. Indeed, even if social preferences would be completely constant over time, we would observe correlation coefficients well below one. Consider a population divided equally between two types, selfish and altruistic ones, where the altruistic type gives to a charity with 20% probability as soon as an opportunity is given, whereas the selfish type never gives. In this case it is easy to see that the correlation coefficient between the contributions to two different charities would be as low as 0.11.[16] The reason for such a low value in this example (even though the actual social preferences are completely stable) is of course that also the altruistic type often gives zero. In this perspective, the obtained correlation coefficients in the present paper are clearly substantial. For example, we consistently find that the correlation coefficient between contributions in 2005 and six years later is as large as 0.2 or larger, despite the fact that the contexts are very different, i.e., the public good in one of these events was the benefit of having a bridge while in the other it was (amplified and shared) money, one of the events was anonymous while the other was not, and one was conducted based on windfall money while the other was not.

Yet we also find, in line with several previous studies, that contributions are highly context dependent. Related to this, we find that some correlation coefficients are substantially larger than others, whether corrected for other explanatory variables or not. Perhaps most strikingly, we obtain a correlation coefficient between voluntary monetary contribution in 2005 and voluntary labor contribution in 2010 in the order of magnitude of 0.4 or larger, despite the fact that almost 50%

[16] From the definition of the correlation coefficient, we have that $\rho = \frac{\mathrm{cov}(x_1, x_2)}{\mathrm{std}(x_1)\mathrm{std}(x_2)} = \frac{E(x_1x_2) - E(x_1)E(x_2)}{\mathrm{std}(x_1)\,\mathrm{std}(x_2)}$, where x_i is the contribution to charity i. From our assumptions, it follows that $E(x_1x_2) = 0.5 \times 0.2 \cdot 0.2 = 0.02$, $E(x_1) = E(x_2) = 0.5 \cdot 0.2 = 0.1$, and $\mathrm{std}(x_1)\,\mathrm{std}(x_2) = (\mathrm{std}(x_1))^2 = \mathrm{var}(x_1) = 0.1 \cdot 0.9^2 + 0.9 \cdot 0.1^2 = 0.09$. Hence, $\rho = \frac{0.02 - 0.01}{0.09} \approx 0.11$.

Table 6
Marginal effects (standard errors) from multivariate regressions; dependent variables are contributions in the three events. Number of observations = 163.

	Only use-the-bridge variables				All variables			
	2005	2009	2010	2011	2005	2009	2010	2011
Use the bridge every day	49.6**	210.0***	1.75***	n.a.	41.8**	213.7***	1.63***	
	(22.2)	(51.1)	(0.66)		(21.3)	(51.1)	(0.62)	
Use the bridge 1–3 times a week	35.1	131.9**	1.63**	n.a.	23.0	111.4*	1.80**	
	(27.4)	(62.6)	(0.79)		(26.4)	(61.9)	(0.76)	
Use the bridge twice a month	−3.1	111.5	−0.17	n.a.	−3.8	120.6	−0.23	
	(23.7)	(48.1)**	(0.75)		(22.6)	(47.7)**	(0.71)	
Use the bridge once a month	22.7	34.7	−0.11	n.a.	14.8	35.1	0.036	
	(19.8)	(41.6)	(0.64)		(18.9)	(41.1)	(0.60)	
Household size					−3.2	16.3	−0.021	3.50
					(4.9)	(11.3)	(0.15)	(6.69)
Age					−0.08	−0.39	−0.051**	1.57
					(0.67)	(1.52)	(0.02)	(0.87)
Male					−12.5	18.6	−0.94	−24.3
					(16.4)	(36.6)	(0.47)**	(22.6)
Education					9.4	27.1	−0.38	−13.2
					(9.7)	(23.9)	(0.30)	(12.8)
Monthly income					−0.26	1.41	0.008	2.69***
					(0.57)	(1.46)	(0.016)	(0.86)
Rice land					4.8**	−0.83	0.160**	9.24***
					(2.3)	(5.48)	(0.07)	(3.21)
Member of the communist party					41.9*	90.0	0.80	12.0
					(24.0)	(63.0)	(0.69)	(34.9)
Member of a local association					18.5	65.1	0.81*	−22.1
					(15.2)	(36.0)	(0.46)	(20.1)
Treatment dummy variables	No	Included	No	No	No	Included	No	No
Experimentalist dummy variables	No	Included	No	Included	No	Included	No	Included

*** indicates significance at the 1% level, ** at the 5% level, * at the 10% level.

contributed nothing in 2005 and over 70% (of the restricted sample that were asked) contributed nothing in 2010. Our conjecture is that this finding may not only reflect stability of social preferences, but also to some extent stability of what may be called sensitivity to social pressure, since none of these events were anonymous. This is an important observation in its own right, and calls for further research. Overall, we conclude that social preferences seem to be quite constant over long periods of time.

4. Conclusions

In this paper we have compared voluntary contributions to a public good, in terms of a bridge in rural Vietnam, for the same complete sample of about 200 households in a village, at four different events spanning over a time period of 6 years. By doing so, we have been able to avoid the potentially confounding factor related to moral licensing and moral cleansing when measuring the extent of pro-social stability over time. Overall, we find substantial and highly significant correlation coefficients, suggesting that pro-social preferences are quite constant over long periods of time.

Although not our main research task, the substantial and positive correlation between the unframed economic experiment and naturally occurring events supports the idea that social preferences obtained in economic experiments have validity also outside the somewhat artificial experimental context. Although our events took place in a village that is typical for this part of the world, it is an open question whether there are large cultural differences in the extent to which social preferences are constant over long periods of time. Previous findings have concluded that there are non-negligible differences in the strengths of social preferences, as measured by economic experiments, in different cultural contexts (e.g., Henrich et al., 2005, 2010). Our conjecture is nevertheless that the extent to which social preferences vary over time does not vary much culturally, although this is of course an open question. For this and other reasons, we encourage further studies in the field in order to test the robustness of our findings.

Acknowledgments

Financial support from Sida to the Environmental Economics Unit at the University of Gothenburg (ID 541000173), EEPSEA (Economy and Environment Program for Southeast Asia) (105899-010), Formas Commons (253-2010-32), the Bank of Sweden Tercentenary Foundation (RS-1319:1), the Swedish Council for Working Life and Social Research (2007-2289), and the Swedish Tax Agency (21-2010-1420) is gratefully acknowledged. We have received valuable comments from two anonymous referees, the editor Tim Cason, Vic Adamowicz, Magnus Johannesson, Peter Martinsson, Adriaan Soetevent, Matthias Sutter, Dale Whittington, Jeffrey Vincent, participants at the Amsterdam Symposium on Behavioral and Experimental Economics, and seminar participants at the University of Gothenburg and the Research Institute of Industrial Economics.

References

Algan, Y., Benkler, Y., Morell, M.F., Hergueux, J., 2013. Cooperation in a peer-production economy: experimental evidence from Wikipedia. Working Paper. Department of Economics, Sciences Po.
Alpizar, F., Carlsson, F., Johansson-Stenman, O., 2008. Anonymity, reciprocity, and conformity: evidence from voluntary contributions to a national park in Costa Rica. J. Public Econ. 92, 1047–1060.
Andreoni, J., 1988. Why free ride? Strategies and learning in public-goods experiments. J. Public Econ. 37 (3), 291–304.
Andreoni, J., 1995. Cooperation in public-goods experiments: kindness of confusion? Am. Econ. Rev. 85, 891–904.
Andreoni, J., Rao, J., 2011. The power of asking: how communication affects selfishness, empathy, and altruism. J. Public Econ. 95, 513–520.
Bardsley, N., Sausgruber, R., 2005. Conformity and reciprocity in public good provision. J. Econ. Psychol. 26, 664–681.
Benz, M., Meier, S., 2008. Do people behave in experiments as in the field? Evidence from donations. Exp. Econ. 11, 268–281.
Blanco, M., Engelmann, D., Normann, H.-T., 2011. A within subject analysis of other regarding preferences. Games Econ. Behav. 72, 321–338.
Brosig, J., Riechmann, T., Weimann, J., 2007. Selfish in the end? An investigation of consistency and stability of individual behavior. FEMM Working Paper No. 05. University Magdeburg.
Carlsmith, M., Gross, A.E., 1969. Some effects of guilt on compliance. J. Pers. Soc. Psychol. 11, 232–239.

Carlsson, F., Johansson-Stenman, O., Nam, P., 2013. Funding a new bridge in rural Vietnam: a field experiment on conditional cooperation and default contributions. Working Paper. Department of Economics, University of Gothenburg.
Cesarini, D., Dawes, C.T., Johannesson, M., Lichtenstein, P., Wallace, B., 2009. Genetic variation in preferences for giving and risk-taking. Q. J. Econ. 124, 809–842.
Cherry, T.L., Kroll, S., Shogren, J.F., 2005. The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. J. Econ. Behav. Organ. 57, 357–365.
Clark, J., 2002. House money effects in public good experiments. Exp. Econ. 5, 223–231.
De Oliveira, A.C.M., Croson, R.T.A., Eckel, C., 2012. The stability of social preferences in a low-income neighborhood. South. Econ. J. 79, 15–45.
Falk, A., Meier, S., Zehnder, C., 2013. Do lab experiments misrepresent social preferences? The case of self-selected student samples. J. Eur. Econ. Assoc. 11, 839–852.
Fehr, E., Gächter, S., 2000. Fairness and retaliation: the economics of reciprocity. J. Econ. Perspect. 14 (3), 159–181.
Fehr, E., Leibbrandt, A., 2011. A field study on cooperativeness and impatience in the Tragedy of the Commons. J. Public Econ. 95, 1144–1155.
Fehr, E., Schmidt, K., 1999. A theory of fairness, competition, and cooperation. Q. J. Econ. 114 (3), 817–868.
Fischbacher, U., Gächter, S., 2010. Social preferences, beliefs and the dynamics of free riding in public good experiments. Am. Econ. Rev. 100, 541–556.
Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Econ. Lett. 71 (3), 397–404.
Frey, B.S., Meier, S., 2004. Social comparisons and pro-social behavior. Testing 'conditional cooperation' in a field experiment. Am. Econ. Rev. 94 (5), 1717–1722.
Gächter, S., Herrmann, B., Thöni, C., 2004. Trust, voluntary cooperation, and, socio-economic background: survey and experimental evidence. J. Econ. Behav. Organ. 55 (4), 505–531.
Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N., Hill, K., Gil-White, F., Gurven, M., Marlowe, F., Patton, J.Q., Tracer, D., 2005. Economic man' in cross-cultural perspective: behavioral experiments in 15 small-scale societies. Behav. Brain Sci. 28, 795–855.
Henrich, J., Ensminger, J., McElreath, R., Barr, C., Barrett, A., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., Ziker, J., 2010. Markets, religion, community size and the evolution of fairness and punishment. Science 327, 1467–1468.
Hoffman, E., McCabe, K., Smith, V., 1996. Social distance and other-regarding behavior in dictator games. Am. Econ. Rev. 86 (3), 653–660.
Isaac, M., Walker, J., Thomas, S., 1984. Divergent evidence on free riding — an experimental examination of possible explanations. Public Choice 43 (2), 113–149.
Jang, K.W., Livesley, J., Vernon, P., 1996. Heritability of the big five personality dimensions and their facets: a twin study. J. Pers. 63, 577–591.
Karlan, D., 2005. Using experimental economics to measure social capital and predict financial decisions. Am. Econ. Rev. 95, 1688–1699.
Karlan, D., List, J.A., 2007. Does price matter in charitable giving? Evidence from a large-scale natural field experiment. Am. Econ. Rev. 97 (5), 1774–1793.
Kroll, S., Cherry, T.L., Shogren, J., 2007. The impact of endowment heterogeneity and origin on contributions in best-shot public good games. Exp. Econ. 10, 411–428.
Laury, S., Taylor, L., 2008. Altruism spillovers: are behaviors in context-free experiments predictive of altruism toward a naturally occurring public good? J. Econ. Behav. Organ. 65, 9–29.
Lazear, E., Malmendier, U., Weber, R., 2012. Sorting in experiments with application to social preferences. Am. Econ. J. Appl. Econ. 4, 136–163.
List, J.A., Lucking-Reiley, D., 2002. The effects of seed money and refunds on charitable giving: experimental evidence from a university capital campaign. J. Polit. Econ. 110 (1), 215–233.
List, J., Berrens, R., Bohara, A., Kerkvilet, J., 2004. Examining the role of social isolation on stated preferences. Am. Econ. Rev. 94, 741–752.
Mazar, N., Zhong, C.-B., 2010. Do green products make us better people? Psychol. Sci. 21, 494–498.
Monin, B., Miller, D.T., 2001. Moral credentials and the expression of prejudice. J. Pers. Soc. Psychol. 81, 33–43.
Shang, J., Croson, R., 2009. Field experiments in charitable contribution: the impact of social influence on the voluntary provision of public goods. Econ. J. 199, 1422–1439.
Soetevent, A.R., 2005. Anonymity in giving in a natural context: an economic field experiment in thirty churches. J. Public Econ. 89, 2301–2323.
Tetlock, P.E., Kristel, O.V., Elson, S.B., Green, M., Lerner, J.S., 2000. The psychology of the unthinkable: taboo trade-offs, forbidden base rates, and heretical counterfactuals. J. Pers. Soc. Psychol. 78, 853–870.
Volk, S., Thöni, C., Rugirok, W., 2012. Temporal stability and psychological foundations of cooperation preferences. J. Econ. Behav. Organ. 81, 664–676.