UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

WORKING PAPERS IN ECONOMICS

No 531

Social preferences are stable over long periods of time

Fredrik Carlsson, Olof Johansson-Stenman, Pham Khanh Nam

April 2012

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

Department of Economics
School of Business, Economics and Law at University of Gothenburg
Vasagatan 1, PO Box 640, SE 405 30 Göteborg, Sweden
+46 31 786 0000, +46 31 786 1326 (fax)
www.handels.gu.se info@handels.gu.se

Social preferences are stable over long periods of time

Fredrik Carlsson, University of Gothenburg, Sweden[A]
Olof Johansson-Stenman, University of Gothenburg, Sweden[B]
Pham Khanh Nam, University of Gothenburg, Sweden and Ho Chi Minh, City University of Economics, Vietnam [C]

Abstract

We measure people's prosocial behavior, in terms of voluntary money and labor time 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 experiments are natural experiment, one is a field experiment and one is a public good experiment. Since the experiments were conducted far apart in time, the potentially confounding effects of moral licensing and moral cleansing are presumably small, if existing at all. Despite large contextual variations, we find a strong positive and statistically significant correlation between voluntary contributions in these experiments, whether correcting for other covariates or not. This suggests that pro-social preferences are fairly stable over long periods of time and contexts.

JEL classification: C93, H41

Keywords: natural field experiment, preference stability, social preferences, moral licensing, moral cleansing.

Acknowledgments: Financial support from Sida to the Environmental Economics Unit at the University of Gothenburg, EEPSEA (Economy and Environment Program for Southeast Asia), and Formas Commons is gratefully acknowledged. We have received valuable comments from Vic Adamowicz, Magnus Johannesson, Peter Martinsson, 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.

[A] Department of Economics, University of Gothenburg, Box 640, 405 30 Gothenburg, Sweden; Ph +46 31 786 41 74; Fax +46 31 786 10 43; E-mail: fredrik.carlsson@economics.gu.se.
[B] Department of Economics, University of Gothenburg, Box 640, 405 30 Gothenburg, Sweden; Ph +46 31 786 25 38 Fax +46 31 773 10 43; E-mail: olof.johansson@economics.gu.se.
[C] Department of Economics, University of Gothenburg, Box 640, 405 30 Gothenburg, Sweden; Ph +46 31 786 47 59; Fax +46 31 786 10 43; E-mail: pham.khanh.nam@economics.gu.se.

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 natural experiments and one field experiment on contributions to a real bridge in rural Vietnam, as well as 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 moment 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, as far as we know, 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 methodologies. 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 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 stable over time. This pattern is similar to the one typically obtained with repeated public good games.[1] Blanco et al. (2011) ran four different experimental games, i.e., dictator, ultimatum, sequential-move prisoners' dilemma, and public good games, and tested whether the Fehr-Schmidt (1999) inequality 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. (2011, in press) 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.

Other studies have also compared contributions in the lab and 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 micro credit 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. Cesarini et al. (2009) used a different approach based on twin studies combined with modified dictator experiments in order 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, is 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

[1] 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).

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.

In summary, there is no consistent pattern from existing studies regarding to what extent social preferences are stable over time. If anything, the existing results seem to indicate that social preferences are relatively unstable over time.

In Section 2, we suggest two related reasons for this pattern among the existing studies, namely moral licensing and moral cleansing, respectively, suggesting that people often seem to have a tendency to compensate for their moral or immoral choices in subsequent actions. We also explain why the potentially confounding effects through moral licensing and moral cleansing do not affect the experimental analysis. Section 3 describes the four experiments and their designs, and provides the corresponding background statistics, while Section 4 presents the results. We find strong positive and statistically significant correlations between voluntary contributions in these experiments, whether correcting for other covariates or not, suggesting that pro-social preferences seem to be quite stable over long periods of time.[2] Section 5 concludes the paper.

2. Dealing with Potential Moral Licensing and Moral Cleansing Effects

One possible explanation for the observed variation in the stability of social preferences over time relates to what psychologists denote *moral licensing* (Monin and Miller, 2001), which suggests that people who have undertaken a praiseworthy act get an implicit license for subsequently conducting a more selfish act. There is a great deal of empirical (in general experimental) support for such licensing effects in the psychology literature. For example, Mazar and Zhong (2010) found that people become less altruistic after purchasing environmentally friendly products than after purchasing conventional products, and Kouchacki (2001) showed that moral licensing effects might also exist at group levels. He found that people were more willing to express prejudiced attitudes when their group members' past behavior had established non-prejudiced credentials.

[2] 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).

Similarly, and symmetrically, there is substantial evidence of *moral cleansing*, which refers to compensatory behavior when people's moral self-worth has been threatened by a morally blameworthy act (e.g., Carlsmith and Gross, 1969; Tetlock et al., 2000). In a recent economics experiments, Gneezy et al. (2011) found, in line with moral cleansing, that people who lied or did not return money they had received by mistake were more likely than others to donate to charity.

Moral licensing and moral cleansing effects taken together seem to suggest that people want to preserve a certain image in the moral domain, an image that, in turn, largely depends on undertaken actions. This is in line with the view of Sachdeva et al. (2009), who also found evidence of both moral licensing and moral cleansing using the same experimental set-up with different treatments that focused explicitly on moral identity. In one treatment, subjects wrote a short story about themselves using nine morally positive trait words (e.g., fair, kind), and in another, subjects used nine morally negative trait words (e.g., selfish, mean). In two other treatments, subjects wrote about someone else using either positive or negative words. The participants were then given a chance to donate part of their compensation to charity. Consistent with moral licensing, subjects in the treatment where they were assigned to write about themselves using positive traits donated the least out of the four treatments, and subjects who wrote about themselves using negative traits donated the most of all, consistent with moral cleansing. There was no difference in donation between the two groups that described someone else, suggesting that the effects are not merely working through priming with moral words. These mechanisms are consistent with a large body of social cognitive research that, according to Dunning (2007), suggests "that people shape their beliefs and judgments of the social world to maintain sacrosanct beliefs of the self as a capable, lovable, and moral individual" (p. 237), and that people's self-worth depends largely on how morally responsible they perceive themselves to be.

Moral licensing and cleansing effects thus constitute potentially confounding effects when testing for stability of social preferences over time. Consider for example a case where a number of people act as senders in two identical dictator experiments (with different receivers). Based on inherent differences in social preferences, one would expect that those who sent more in the first round would also send more in the second. Yet, based on moral licensing or moral cleansing (depending on the reference points for bad versus good actions), one would expect that an individual who sent more in the first round would for this reason send less in the second. One way around this confounding effect would be to space the tests far apart in time in order to eliminate any moral licensing and moral cleansing effects. This is

the strategy obtained in the present study. Another obvious advantage of the large time span is that we can test whether the underlying preferences are the same for long periods of time.

3. The Four Experiments

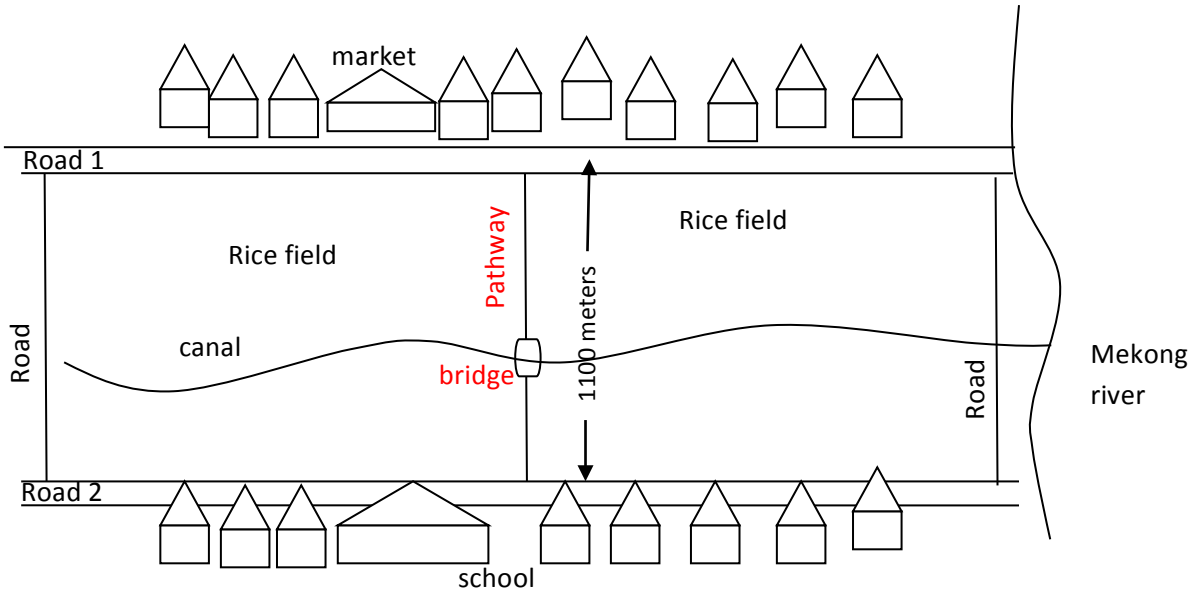
We use observations on subjects' pro-social behavior in four related events, or experiments, in 2005, 2009, 2010, and 2011, i.e., with a rather long time period in between and spanning over about six years. Two of the events (in 2005 and 2010) are naturally occurring ones where we simply observed the behavior, and may hence be classified as natural experiments. The two others (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 linked to the bridge at all.

All four experiments focus on voluntary contribution mechanisms, and although there are a number of contextual differences, in each experiment we observe the behavior of the same (approximately 200) subjects, representing all households in the village. The experiments were undertaken in the Giong Trom hamlet,[3] in the Mekong river delta of Vietnam, where 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. 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. Small-scale infrastructures within a hamlet are considered to be the responsibility of the hamlet.

3.1 The bridge and the experiments

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 1,200 meters from the bridge's pathway; see the following map.

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



The first three experiments, to be described next, concern funding of a bridge for the hamlet.

3.2 The 2005 experiment

In the first experiment 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.[4] 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 its shape was as shown in the picture below.

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

As can be seen, the wooden bridge was highly degraded, and could obviously not be used for tractors and motorbikes.

3.3 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; for a detailed description of the experiment and the results, see Carlsson et al. (2010). 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 experiment involved five treatments in which one treatment served as a reference case and the others involved reference contributions and default options. 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 fulfill the threshold, the new bridge was built in early 2010; see the picture below.

3.4 The 2010 experiment

The experiment in 2009 thus resulted in the construction of the bridge in 2010 since the total contributions were higher than the threshold. In preparing for 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. The construction work required joint efforts in a short time period. According to the construction process, a number of particular days were scheduled for the joint work. Two persons from the hamlet council visited the households in the hamlet to invite villagers to contribute labor to build the bridge. Hence, an important difference compared to the previous two experiments is that instead of being asked for monetary contributions, they were asked for labor contributions.

However, not all households were asked to make contributions, since some households were not expected to be able to contribute any labor at all, mainly because the household members were too old. In total 19 percent of the households were not asked to make any labor contribution.[5] At this time, households were not told anything about what others were contributing, and there was obviously 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 experiment is the actual amount of labor contributions, and not what they promised when asked to contribute.

3.5 The 2011 experiment

While the bridge is presumably useful to all households, the usefulness varies with, e.g., 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 related to the bridge. Thus, the experiment conducted in 2011 was not directly related to the bridge.

We chose to conduct a public good game, much like a standard 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,

[5] 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.

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.

3.6 Summary of experimental designs and household characteristics

As mentioned, the first three experiments related to the bridge in the hamlet. The first experiment 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 experiment concerned voluntary labor contributions to build the road to new concrete bridge in 2010. The fourth experiment was instead a public good experiment that was not related to the use of the bridge. The settings of the four experiments are summarized in Table 1.

Table 1. Characteristics of the four experiments

Characteristics	Experiment 1	Experiment 2	Experiment 3	Experiment 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	Fundraising campaign	Public good game
Windfall money	No	Yes	No	Yes
Contribution range	[0, .) thousand dongs	[0, 400] thousand dongs	[0, 3.5] labor days	[0, 200] thousand dongs
Organizer	Local government	Outside NGO	Local government	University
Reference contribution	Yes	Yes in some treatments	No	No

While we designed only two of the four 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, of the households. The total

number of households in the village is 200. However, in the last experiment, four of the households were unable to participate.[6]

Table 2. Household characteristics

Variables	Definition	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 every day	= 1 if use bridge every day	0.19	
Use the bridge 1-3 times a week*	=1 if about 1-3 times a week	0.10	
Use the bridge twice a month	= 1 if about 2 times a month	0.17	
Use the bridge once a month	= 1 if 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	
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 hectare)	4.69	3.13

* The options for the question regarding the 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 chose options 2 and 3, we merged them in the descriptive statistics and in the analysis.

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. The households in the study are thus poor, and the average education level

[6] 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.

is very low. The average size of land on which a family is currently cultivating rice is also rather small, approximately half a hectare.

4. Results

4.1 Average contributions in the four experiments

Before looking at the correlations between the contributions, let us briefly look in Table 3 at the average contributions in each of the experiment, based on 196 households. Since not all households were asked to contribute labor in the experiment in 2010, we also present the statistics of contributions both for the whole sample and the restricted sample of households that had the possibility to contribute in 2010.

Table 3. Descriptive statistics of contribution variables in three experiments

Experiment	Mean	Std. dev.	Share zero	Min	Max
2005 (thousand dong)	40.10	56.15	0.46	0	300
2009 (thousand dong)	270.26	127.95	0.02	0	400
2010 (labor days, whole sample)	0.41	0.86	0.76	0	3.5
2010 (labor days, restricted sample)	0.50	0.92	0.71	0	3.5
2011 (thousand dong)	125.92	68.27	0.05	0	200

Comparing the 2005 and 2009 contributions, which were both in terms of monetary contributions to a new bridge, there are strikingly large differences. The average contribution in 2009 was almost seven times as large as in 2005, 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, two stand out clearly: First, contrary to in 2005, 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). Second, and again contrary to the 2005 experiment, the experiment in 2009 contained a windfall endowment provided by the NGO involved in the experiment.[7] Moving next to the 2010 experiment, we can observe that even fewer chose to contribute than in 2005. In 2010, the average contribution of labor was 0.4 labor days per household, which corresponds to about 32,000 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, and the average contribution is

[7] 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-gains effect on contributions, while Kroll et al. (2007) found significant differences in a public good experiment with heterogeneous endowments.

substantial. Interestingly, the mean contribution as a fraction of the maximum contribution is very similar as in the 2009 experiment. These two experiments share the features that they are concerned with voluntary financial contributions to a public good and also that the contributions are based on windfall endowments, where the latter presumably is an important factor in explaining the higher contribution rates. Yet, we are not primarily interested in to what extent the contribution levels are the same across the decisions, but to what extent decisions are correlated, i.e., whether or not those who contribute more in one experiment also contribute more in another.

4.2 Raw contribution correlations between the experiments

As described above, we observe the contributions in each of the experiments at the household level. As a first step, we therefore analyze the simple pairwise correlations between the four experiments. Remember that we have three observations of contributions to the bridge from the same household, 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 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. In Table 4, the pair-wise correlation coefficients are presented.

Table 4. Correlation coefficients, contributions in the experiments

	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.

Despite the large differences in contribution levels between the experiments, including in the fraction that did not contribute anything, the correlation coefficients between the four

experiments are substantial and in most cases highly significant. For the experiments related to the bridge, the largest correlation coefficients are found between the 2005 and 2010 experiments. This may seem surprising since the 2005 experiment concerns monetary contributions, while the 2010 experiment 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. However, we see two main explanations: First, in both of these experiments 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 experiments 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 experiments, implying a positive effect on the correlation coefficient.

The correlation coefficients between the experiment in 2011 and the other three experiments are also substantial, with the exception of the experiment 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 differences in experimental set-up. Together, this clearly shows (1) that the strong correlations between the experiments cannot only be due to the fact that they concern contributions to a similar good, i.e., the bridge, and (2) that there is clear support for the idea that social preferences reflect traits that to a large extent are constant over time and domains.

4.3 Econometric analysis

While the strong positive correlation coefficients between contributions in the first three experiments (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 behind these positive correlations. For example, if the households who use the bridge the most are also willing to contribute the most (e.g., for selfish reasons), we should obtain a positive correlation between contributions in the three

first experiments even if there are actually no differences between the households in terms of underlying social preferences.

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 in order to correct 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 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 experiments in 2005 and 2010. Using a multivariate model, we estimate the correlation coefficients of the error terms for each experiment. 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 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 experiments, since these variables presumably vary across the households and at the same time are independent of underlying differences in social preferences, and for the second and fourth experiment 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 or income, that are correlated with true underlying social preferences. For example, suppose that all variation in social preferences is determined by gender. If we then correct for gender in the regressions, we will find that there is no stability of social preferences over time, even though there may perfectly well exist a certain degree of stability in reality (through gender). 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 that were not asked to contribute.

We also use the full data set for the pairwise correlations that do not include the 2010 experiment. Yet, as can be observed, the results turned out fairly similar.[8] The estimated correlation coefficients for our three sets of multivariate regressions are presented in Table 5, for each separate experiment. In the appendix, we also report the estimated coefficient for the covariates.[9] Starting with the first three experiments, we can observe that the pairwise correlation coefficients are consistently positive, substantial, and statistically significant. Consequently, even when controlling for a number of observable differences among households and the treatment effects, there are strong correlations in behavior between the three experiments. 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 experiments, probably largely due to the fact that these experiments were not anonymous as discussed previously.

[8] 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.

[9] Few of the household characteristics have a significant impact on the contributions in any of the experiments. Furthermore, there is no consistent pattern across the three experiments. The contributions in 2005 are positively correlated with the size of the land and with whether any household member is a member of the communist party. The contributions in 2009 are only positively correlated with the use of the bridge. In addition, some of the treatment dummy variables, not reported here, are statistically significant. The contributions in 2010 are positively correlated with the size of the land, membership in local associations, and use of the bridge, and are lower if the age of the household head is higher. The contributions in 2011 are positively correlated with income and the size of the land.

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 experiments.

	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.

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 public good field 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 experiments concerned the same underlying good, a bridge, that explains the significant correlation between the contribution decisions.

4.4 Are the obtained correlations 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. Yet, 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 only, 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 why the correlation coefficient between the contributions to two different charities would be as low as 0.11.[10] The reason for this relatively low value is of course that also the altruistic type often gives zero. In this perspective, the obtained correlation coefficients here 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 experiments was the benefit of having a bridge while in the other it was (amplified and shared) money, one of the experiments 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

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

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 contribution in labor time in 2010 in the order of magnitude of 0.4 or larger, despite the fact that almost 50% 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 of social pressure, since none of these experiments 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.

5. Conclusion

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, spanning over a time period of 6 years and using a combination of two natural experiments and two field experiments. 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 artefactual field experiment and the other experiments also contributes to the literature on external validity of laboratory methods. Our results thus support the idea that social preferences obtained in economic experiments have validity also outside the somewhat artificial experimental context.

Although our experiments were conducted 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 parts of the world. Our conjecture is nevertheless that the extent to which social preferences vary over time is fairly constant, yet, again, this is an open question. For this and other reasons, we encourage further experimental studies in the field in order to test the robustness of our findings.

References

Alpizar F., Carlsson F., Johansson-Stenman O. 2008. Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica, Journal of Public Economics 92, 1047-1060.

Andreoni, J. 1988. "Why Free Ride - Strategies and Learning in Public-Goods Experiments." Journal of Public Economics, 37(3): 291-304.

Andreoni, J. 1995. Cooperation in public-goods experiments: Kindness of confusion?, American Economic Review 85, 891-904.

Bardsley, N. and R. Sausgruber (2005) Conformity and reciprocity in public good provision. Journal of Economic Psychology 26, 664–681.

Benz, M. and S. Meier. 2008. Do people behave in experiments as in the field? – Evidence from donations. Experimental Economics. 11: 268 – 281.

Blanco, M., D. Engelmann, and H.-T. Normann (2011), A Within Subject Analysis of Other Regarding Preferences, Games and Economic Behavior 72, 321-338.

Brosig, J., T. Riechmann and J. Weimann. 2007. Selfish in the end? An investigation of consistency and stability of individual behavior. FEMM Working Paper No.05. University Magdeburg.

Carlsmith, M. and Gross, A. E. (1969). “Some Effects of Guilt on Compliance,” Journal of Personality and Social Psychology 11: 232-239.

Carlsson, F., O. Johansson-Stenman and P. Nam. 2010. 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 CT, Johannesson M, Lichtenstein P, Wallace B. 2009. Genetic variation in preferences for giving and risk-taking. Quarterly Journal of Economics 124: 809-842.

Cherry, T.L., S. Kroll and J.F. Shogren, 2005, The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab, Journal of Economic Behavior and Organization 57, 357–365.

Clark, J., 2002, House money effects in public good experiments, Experimental Economics 5, 223-231.

De Oliveira, A.C.M., Croson, R.T.A. and Eckel, C., 2011. The Giving Type: Identifying Donors. Journal of Public Economics 95, 428-435.

De Oliveira, A.C.M., Croson, R.T.A. and Eckel, C., in press. The Stability of Social Preferences in a Low-Income Neighborhood, Southern Economic Journal.

Dunning, D. (2007). Self-image motives and consumer behavior: How sacrosanct self-beliefs sway preferences in the marketplace. Journal of Consumer Psychology, 17, 237–249.

Fehr, E., and S. Gächter (2000) "Fairness and Retaliation: The Economics of Reciprocity", Journal of Economic Perspectives, Vol. 14, No. 3, pp.159-181.

Fehr, E. and Schmidt, K., 1999. A theory of fairness, competition, and cooperation. Quarterly Journal of Economics 114 (3), 817-868.

Fischbacher, U. and S. Gächter. 2010. Social preferences, beliefs and the dynamics of free riding in public good experiments. American Economic Review 100: 541-556.

Fischbacher, U., S. Gächter and E. Fehr 2001. Are People Conditionally Cooperative? Evidence from a Public Goods Experiment. Economics Letters, 71(3): 397-404.

Frey, Bruno S., and Stephan Meier. 2004. Social Comparisons and Pro-Social Behavior. Testing 'Conditional Cooperation' in a Field Experiment. American Economic Review, 94(5): 1717-22.

Gneezy, U., A. Imas and K. Madarasz (2011), Conscience Accounting: Emotional Dynamics and Social Behavior, working paper.

Hoffman, E., K. McCabe and V. Smith (1996) "Social Distance and Other-Regarding Behavior in Dictator Games", American Economic Review 86(3): 653-660.

Isaac, M., J. Walker and S.Thomas 1984. Divergent Evidence on Free Riding - an Experimental Examination of Possible Explanations. Public Choice, 43(2): 113-49.

Karlan, D. 2005. Using experimental economics to measure social capital and predict financial decisions. American Economic Review 95: 1688-1699.

Karlan, D. and J. A. List, 2007. Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment. American Economic Review, 97(5): 1774-1793.

Kouchaki, M. (2011), Vicarious moral licensing: The influence of others' past moral actions on moral behavior. Journal of Personality and Social Psychology, 101(4), 702-715.

Kroll, S., T.L. Cherry and J. Shogren, 2007, The impact of endowment heterogeneity and origin on contributions in best-shot public good games, Experimental Economics 10, 411-428.

Lazear, E., Malmendier, U., and Weber, R. (2012). Sorting in Experiments with Application to Social Preferences. American Economic Journal: Applied Economics 4, 136-163.

Laury, S. and L. Taylor. 2008. Altruism spillovers: Are behaviors in context-free experiments predictive of altruism toward a naturally occurring public good? Journal of Economic Behavior and Organization 65: 9-29.

List, J. A. and D. Lucking-Reiley 2002. The Effects of Seed Money and Refunds on Charitable Giving: Experimental Evidence from a University Capital Campaign. Journal of Political Economy, 110(1), 215-233

List, J., R. Berrens, A. Bohara and J. Kerkvilet (2004) Examining the role of social isolation on stated preferences. American Economic Review 94, 741–752.

Mazar, N. and C-B. Zhong. 2010. Do green products make us better people? Psychological Science 21: 494-498.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. Journal of Personality and Social Psychology, 81, 33–43.

Sachdeva, S. Iliev, R. and D.L. Medin (2009). Sinning Saints and Saintly Sinners. The Paradox of Moral Self-Regulation, Psychological Science 20: 523-528.

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. Journal of Public Economics 89, 2301–2323.

Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. and Lerner, J. S. (2000). The Psychology of the Unthinkable: Taboo Trade-Offs, Forbidden Base Rates, and Heretical Counterfactuals. Journal of Personality and Social Psychology, 78, 853–870.

Appendix

Table A1. Marginal effects (standard errors) from multivariate regressions; dependent variables are contributions in the three experiments. 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)
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