Supporting Information for

"Fragility of the provision of local public goods to private and collective risks"

Juan-Camilo Cárdenas, Marco A. Janssen*, Manita Ale, Ram Bastakoti, Adriana M. Bernal, Juthathip Chalermphol, Yazhen Gong, Hoon C. Shin, Ganesh Shivakoti, Yibo Wang, John M. Anderies

* Corresponding author: Marco.Janssen@asu.edu

This file includes:

- I. Experimental procedures
- II. Individual and community data
- III. Context and sampling of communities
- IV. Theoretical predictions on risk and cooperation
- V. Robustness checks on results

I. EXPERIMENTAL PROCEDURES

Experimental Methods

In this section we detail our experimental procedures and protocols for work reported on in this publication. The complete set of procedures, protocols and scripts themselves, as well as data collection sheets are available at osf.io/kgz2r.

Game procedures and protocols

General goals

The original experimental design was developed in English, translated by the project members into Spanish, Thai, Nepalese and Chinese. Subsequently, the experimental protocol was back translated by colleagues who were not involved in the project. The protocols were approved by the IRB at Arizona State University.

The experiment was executed in each country by local research groups who performed the experiment in their native language. Each research group was trained in experimental methods during a workshop at Asian Institute of Technology in 2012 where Cardenas and Janssen were the main instructors. During that workshop an initial design was pretested with rice growing farmers near Bangkok. In 2013, colleagues from Thailand, Nepal and China visited the group of Cardenas in Colombia who started the experiments to undergo additional training directly in the field to ensure the protocol would be executed in each country in the same way.

In the design of the experiment, survey and interview questions, we took into account that we would have one day to run the experiment and collect the data in each community, and that in some communities we will have to work with participants with low literacy levels.

At the end of a session, each participant would have earned around 1-2 days wage of labor at the local wage rate for unskilled labor. All participants also received a show-up fee to cover for transportation costs and equivalent to approximately one third of a day's wage. Participants signed a consent form in which it was explained that their participation was voluntary, that they would be paid upon completion of the tasks in the experiment, and that their individual information was to be kept confidential by the experimenter.

Public goods game description:

The framing used for the experiment is within the context of an irrigation system that provides water for the rice farming, and that is maintained by the contributions of the water user farmers. The game is presented in the following manner:

> *"**This is a game about irrigation:***
> *In this exercise you are part of a group of farmers that use an irrigation system which provides water for your crops. Since the crop requires water, your income in this game depends on the availability of water for its irrigation. In order to irrigate your crops in the group of farmers you belong to, there is an irrigation system that benefits the community, such as this one in the board.*

[At this point the experimenter shows an illustration of the irrigation canal and users, as shown in the picture in Figure S1]

Figure S1. Experimenter shows an illustration of the irrigation canal.

> *"However, the water you get from the canal depends on how well maintained is the irrigation system. During this game you will make several decisions in different rounds with respect to your crops and the irrigation system maintenance."*

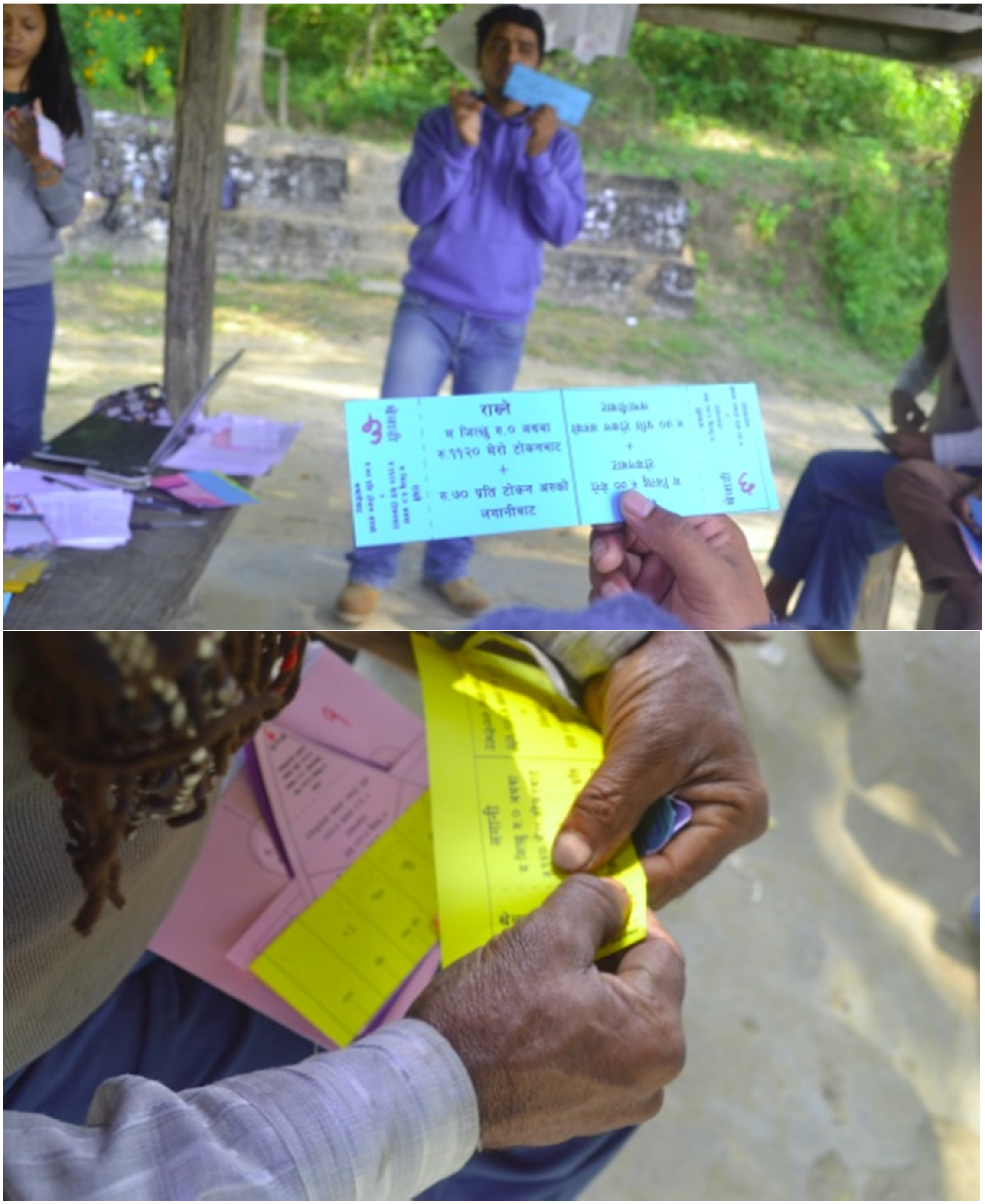

Figure S2. Slips of decision cards, using colors to facilitate comprehension of the task.

We use a linear public good game where in the first round we present participants with the decision to either invest or not to invest in the public fund. The experimental instructions are read out loud and with examples to test the participants' understanding of the instructions. We explain here the experiment in US dollar terms, but each experiment was performed in local currencies, guaranteeing that we paid amounts of money that had the same purchasing power in the rural settings of our participants (See Table S4 for the rate of conversion used based on prices for regular consumption goods).

In our experiment each of the players had to choose between keeping an endowment (one token) with a private return of $8, or investing it in a group fund that provided a public good which transferred to each player $1 for each token invested by the group. To make such decisions, the players received slips of papers as shown in the pictures in Figure S2. Each color represented a different round of the game, and to mark the choice, they ripped either end of the slip, depending on whether they wanted to invest in the public fund or keep the token for themselves.

Notice, for groups larger than 8 people, and because the ratio of the marginal return on the public good to the return of the private option is 1/8=0.125, we have a clear social dilemma to be solved by these individuals. For our target group size of 20 players, full cooperation would yield $20 for each player, whereas universal defection would leave each player with only $8. However, regardless of the number of cooperators in the group, the dominant strategy all players is to keep the token and earn those extra $8 on top of the amount received form the public fund.

There is no communication possible between the participants and each makes their private decision simultaneously. Such decisions are kept privately and confidentially by the experimenter, according to the signed consent form. The results of each round, including their payments, were presented only after all the experimental tasks were completed, reducing the learning and reciprocity effects over rounds. We also asked participants to guess how many people would invest their token in the group fund for each of the three rounds. For this, the participants received each another slip of paper (Shown in figure S3 and in the picture in figure S4) where they will mark their best guess, from zero contributions in the left option to full cooperation in the right.

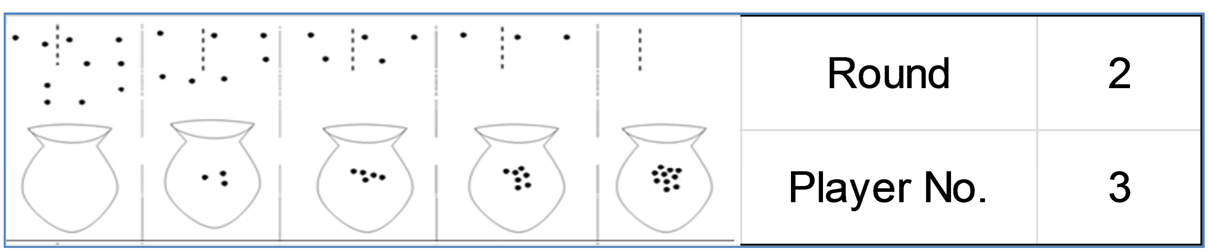

Figure S3. Example of the slip of paper to indicate the estimate of the amount invested in the public good by the group.

Figure S4. Participant indicate his/her expectation of the level of group investment.

After the first round, we conducted our next two rounds under risk. We alternate among the communities as to whether round 2 or round 3 is the one with risk to the return from the private endowment or public endowment (PrivateRisk=blue round and CollectiveRisk=yellow round). We do this to have a clean comparison between subjects, and make sure we do not have order or learning effects of concern. In fact, we do not find any order effects of concern in the data. A Two-sample Wilcoxon rank-sum (Mann-Whitney) test between Private vs Collective Risk for the second round yields a p-value=0.0000, with the collective risk showing a 12% lower rate of cooperation than under private risk. If we compare the Private vs Collective Risk treatments in round 3, we also obtained the same direction and statistically significant difference, with a 5% lower rate of cooperation for the collective risk treatment (p-value= 0.0176. This makes it possible to conduct a clean between-subjects comparison in round 2 of the procedure, and against the baseline data from round 1.

Blue round (rounds 2 or 3, risk affects return from private asset):
The game in this round works very similar to round 1 except for one condition. This time the amount one earns from keeping the token is not $8. It can be higher or it can be lower than $8. After the decisions are made the experimenters pull a ball from a bag that has 5 red balls and 5 blue balls. If we draw a red ball one will earn $16 from the token kept, if one gets a blue ball one will get nothing ($0) from that token. Regardless of investing or keeping the token, one will still earn $1 per token in the group fund. Those groups, randomly selected to play this task as their second round, then proceeded to the “Yellow round” in round 3.

Yellow round (rounds 3 or 2 respectively, risk affects return on public good):
The game in the yellow round is similar to round 1, except for one difference. This time, the amount that a person earns from keeping the token remains at $8, but the amount that one earns from each token in the group fund will change. It can be higher or it can be lower than $1 depending on the ball drawn from the bag, with a 50% chance that everyone in the group will earn $2 per token in the group fund, or nothing ($0) per token in the group fund. Those groups,

randomly selected to play this task as their second round, then proceeded to the "Blue round" in round 3.

We announced to them, in advance, that at the end of the experiment one of these rounds had a random chance of being chosen for payments to avoid participants averaging their earnings across rounds. Results for that round in terms of the group investment and of the outcome of high/low returns from the group fund or private endowment, as appropriate were only announced once all decisions were made in the experiment to avoid retaliation or learning effects over rounds.

Measuring risk aversion

Also, to provide additional context for the interpretation of our data we asked participants to choose among a set of six lotteries, using methods (1,2) to elicit their individual risk preferences. The data from these individual preferences were used as controls in our regression analysis. The following figure is an example of the lottery choices used in China.

The payoffs for each 50–50 lottery were chosen so that the expected payoff of each lottery increases as one moves clockwise (from RMB 13 yuan to RMB 18 yuan versus RMB 10 yuan), but so does the variance of the payoffs. The payoffs in each circle used in other countries were adjusted according to purchasing power in each country in 2013.

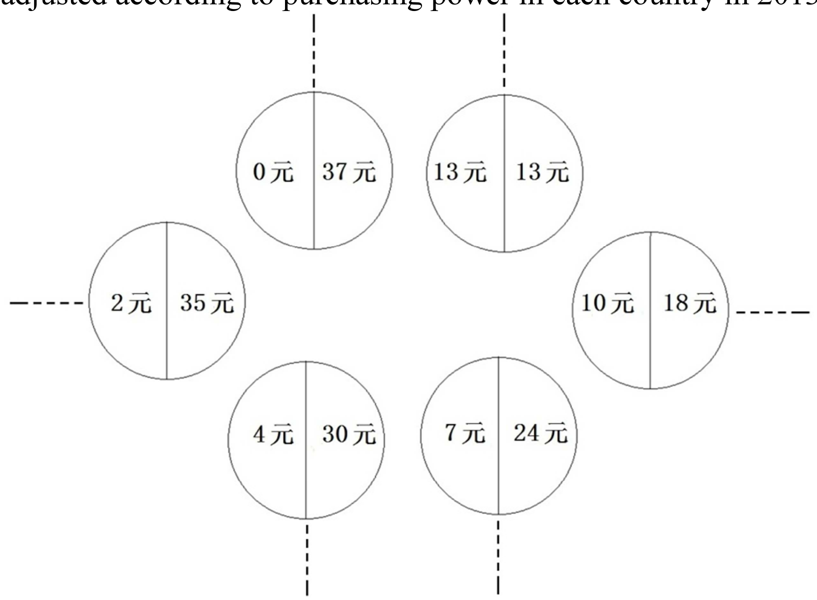

Figure S5: Lottery options, with a 50/50 probability, to elicit risk preferences from the participants (1,2).

To characterize the risk attitudes, a constant relative risk aversion (CRRA) utility function, $U(x)=\frac{x^{1-r}}{1-r}$, is used. CRRA utility has been used to describe risk preferences in a number of experimental studies (2,3,4). At any switching point, people should be indifferent between two neighboring lotteries.

Conversion rates and payoffs equivalents:
In order to make comparable payments in all four countries we estimated in each of the rural areas of these countries the prices for a basket of goods and the value of the minimum wage available with such prices we estimated a conversion rate so that the earnings in the game had the same purchasing power across all countries. The following Table S1 has these prices, at the USD conversion rate at the time of the field work (Summer and Fall, 2015).

	Colombia		**Thailand**		**Nepal**		**China**	
ITEMS	USD	Pesos	USD	Bahts	USD	Rupee	USD	Yuan
1 day wage	14	25000	10	300	6.25	500	N/A	N/A
lunch in the field	2	3776	1.66		2		2.43	
3 kg rice bag	3.7	6985.6	3.5		2.5		2.14	
1 kg sugar bag	1.3	2454.4	0.8		1		1.62	
1/2 liter cooking oil	2.3	4342.4	0.92		1		0.97	
1 kg salt	0.5	944	0.67		0.3		0.65	
Basket	9.8	18502.4	7.55		6.8		7.80	

Table S1. Purchasing power from payments in the games.

II. INDIVIDUAL AND COMMUNITY DATA

In each community we collected data on individual and community level economic, demographic and other contextual variables using standardized collection protocols and forms.

Individual Survey
Upon completion of the experimental procedure (public goods game and risk preferences game) each participant filled out a survey with questions related to social-demographic variables, their professional and agricultural experience and the social fabric of the community.

Age, sex and education
Each participant provided information on their age (years of age) and sex (male/female). We also asked participants to indicate their level of education for which we ask the highest grade they have completed in school. The categories, that allowed us to compare across all four countries were (0) none; (1) Adult literacy; (2) Some primary school; (3) Primary School completed; (4) Secondary School completed; (5) Technical; and (6) University. To control for different educational systems among the countries, we normalized the educational data for each country level by subtracting the mean education response of that country and dividing by the standard deviation.

Household size, income and land size
For each participant we asked the number of people (adults and children) who live in the house during the last year and for at least six months. The response to this question was used as the household size. To derive an estimate for the annual income we asked participants to estimate the annual or average monthly income for the following sources: agriculture, business, services, remittances, pensions, off-farm wage earnings and other income. All the responses were translated into dollar amounts and added up to derive an estimate of the annual income. In our

analysis we normalized income data for each country level by subtracting the mean income of that country and dividing by the standard deviation. Each participant was asked whether they owned the land they live on and the size. The responses were all translated in terms of acres of land. Also the land size data was normalized for each country between 0 and 1 (largest land size in the sample of the country).

Trust

Building a reliable measurement of trust is a key control to include for our behavioral measurements regarding the social fabric of the community. However, the more general questions used in surveys like the World Values Survey may hinder some of the details of what happens in the internal dynamics of a community. Further, stated behavior in surveys and revealed actions in experiments can also be complemented as ways of explaining prosocial behavior in individuals and communities (See (5) for a large study using these two methods). In order to capture the essence of inter-personal trust in terms of a more day-to-day realm, we use more familiar situations our participants may have faced.

To derive an estimate of trust in other community members we ask whether participants agree or disagree with a number of statements. The responses can be (1) Disagree Strongly; (2) Disagree; (3) Agree or (4) Agree Strongly.

We use the following five questions.

1. Most people in this community are basically honest and can be trusted.
2. In this community, one had to be alert or someone is likely to take advantage of you.
3. If a mother in this community has an emergency and needs to leave her baby for the day, she will easily find someone in this community she can trust with her baby.
4. If someone loses a pig, goat or chicken he or she will easily find others in this community to help to seek and find it.
5. If a neighbor in this community lends some money to another neighbor, it is very likely that the lender gets her money back

To derive a value between 0 and 1, we use the following formula to scale the responses, where X_i is the response to question i: $(X_i – X_2 + X_3 + X_4 + X_5)/20$

Interview

To derive information about the community, their governance structure and economy, interviews were held with a group of community leaders, typically members of the water governance board (such as the Water Users Association in Thailand and Nepal)). Such interviews were conducted after the experiments were conducted. Questions included in this interview (as used in our analysis) were:

1. What is the number of households in the community?
2. What is the percentage of community members working and living outside the community (say, in a nearby city)?
3. What percentage of total crop cultivation area in this irrigation system's command area is grown for such market-targeted production?
4. What percentage of the food consumed by community members is produced in the community?

III. CONTEXT AND SAMPLING OF COMMUNITIES

Our choice for the sample of these four countries and the selection of rice as the common crop across all countries was not accidental. We chose a crop that was sufficiently important in terms of area planted and food security, and that was common in all countries where our field and research teams had experience. Secondly, we wanted a crop that was highly dependent on water as an input for production. Having a common crop would maintain the comparability in terms of dependence on water and water irrigation and with similar production functions in terms of the inputs and technology.

Rice is a crop of major relevance for food security worldwide, and is grown in most regions where concerns over water and climate change are on the rise. As seen in the figure below, the relevance of rice among the top crops worldwide is clear if seen in terms of area cultivated, similar to maize and below wheat as the top crop in terms of acreage.

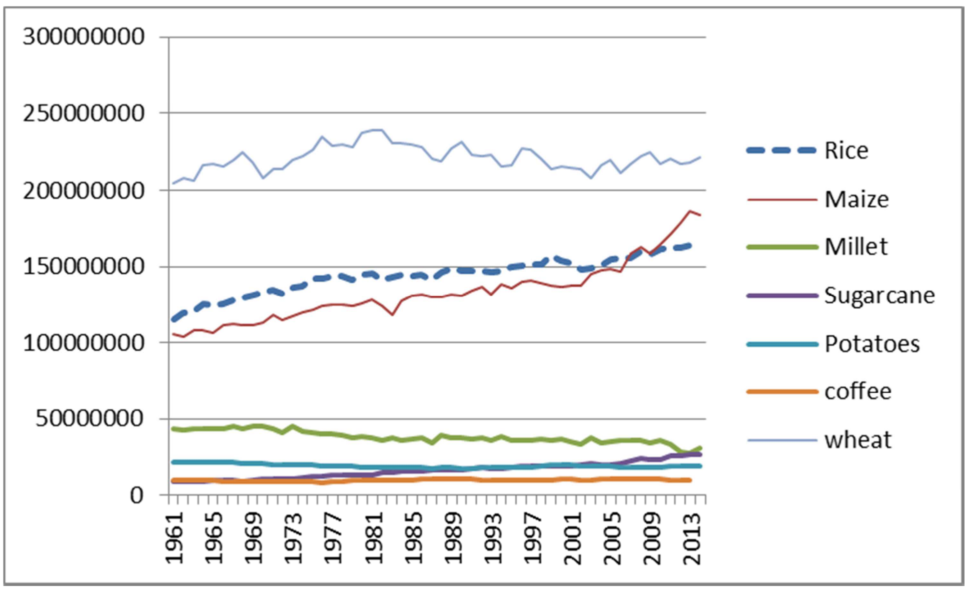

Figure S6. Area cultivated for the top crops in the world. Source FAOSTAT 2016.

Also, we wanted to have variation across countries, with those where rice is a major crop in terms of area, yields and population involved, and others with less prevalence. As shown by (6), China and Thailand are among the top ten exporters of rice, whereas its relevance in terms of markets integration and consumption are lower for our other two countries in the study. Nepal derives an important part of its diet and agriculture from rice farming, mostly for self-consumption, and Colombia with a rather small rice farming sector and often a net importer.

In terms of variability regarding water stress levels and the relevance of these crops, (7) show that basins such as the Mekong and Brahmaputra basins show high variations in climate and low storage capacity, precisely where paddy rice is a major crop, again, highly dependent on water and subject to higher probabilities of stress if densification and economic volatility continues to grow.

Sampling
In each of these four countries we identified regions where rice production was relevant, and within those regions we visited 30 communities in each of these countries, except for Colombia where we could only reach 28.

Around 20 participants were recruited in each of our 118 communities. Participants needed to be 18 years or older and only one person per household was allowed to participate. Local teams under the main researchers and co-authors of the article worked with community leaders to identify a random sample in the area, and that were involved in rice production. When households records were available, random samples of 20 participants were drawn, otherwise the researchers randomly identified 20 houses to recruit one participant in each. Participation was voluntary in all cases, using a consent form, signed by the researcher and describing also the characteristics of the exercise, and the confidentiality of the data.

Often the participants were members of the irrigation or farmers' organization, and recruitment was then organized via such an organization. In occasions more than 20 people signed up and thus we accepted them. The theoretical predictions of a Nash equilibrium with zero cooperation for individuals that maximize their own material payoffs is not altered by these small variations in group size, as long as the group size is greater than eight participants in the game.

Experiments were performed during months where there was a higher likelihood that farmers were available (different than planting or harvesting seasons). Some individuals from these households were out of the community to earn additional income.

In each of the community the experimental component of the field work was conducted entirely within the first day the team stayed at that community.

Informed consent
All participants signed a consent form before they could participate in the experiment.

Selection of Regions and Communities
Experiments were scheduled to be performed in 30 communities in each participating country (China, Nepal, Thailand, and Colombia). The eligibility of communities included some form of water management system, and being a rice-growing area. The regions and communities selected are described next and shown in Figure S7, as well as in the main article.

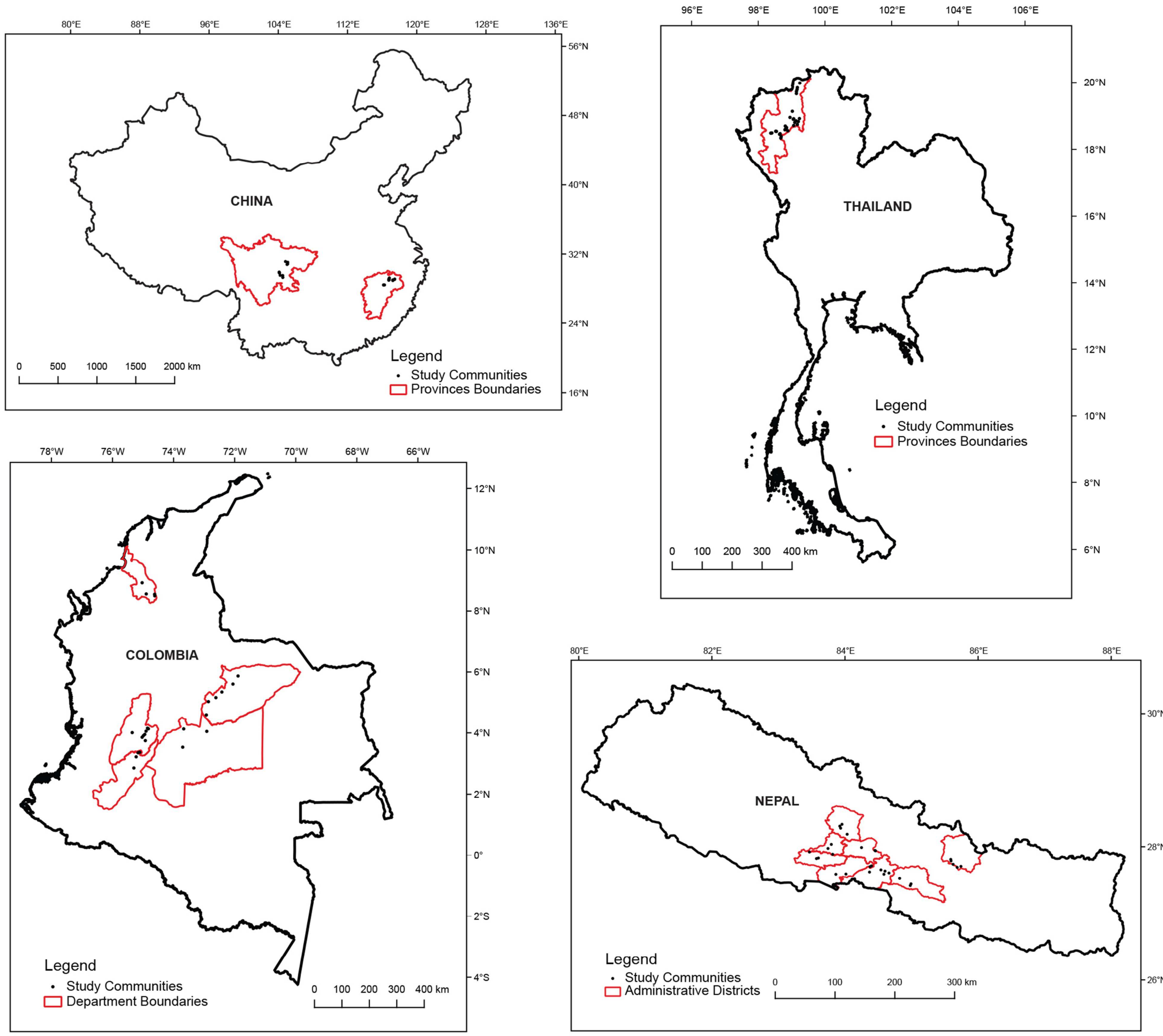

Figure S7. The following maps show the regions (in blue circles) where the samples of communities were drawn for our experiments (Left: Colombia, Right: Nepal, Thailand, China)

About the context in each country and regions.

In **Nepal** we sampled 8 districts from various parts of the country. All the study districts are characterized by the presence of small scale irrigation systems those supply irrigation water to rice crop in monsoon season. The study district covered three different physio-graphic zones and exposures to natural disturbances such as drought and flooding. The situation varied across the three physio-graphic zones. Sindhupalchowk district was chosen from Bagmati River Basin that is in mid-mountain zone. Rice was main crop in monsoon season whereas other seasons dominated by maize and other crops. Five irrigation systems included as sample. All selected irrigation systems were small farmer managed systems. Drought and flash floods affected canal infrastructure regularly. From mid-hills zone Palpa, Syangja, Kaski and Tanahu districts of the Gandaki River Basin were included in the study. Rice was the main crop in the monsoon season, while wheat dominated in winter and maize in the pre-monsoon season. In addition, commercial vegetables were grown in some parts of those districts. Three irrigation systems from Palpa, three systems from Syangja, four systems from Kaski and three systems from Tanahu were included in the sample. Small farmer managed irrigation systems dominated in these districts. Flash flooding often damages the irrigation infrastructure, whereas drought was also common problem. Furthermore, in the Terai Plains and Siwalik zone, Chitwan and Makwanpur districts in East Rapti River Basin and Nawalparasi district in Gandaki River Basin were chosen for the study. Rice is the main crop in monsoon, wheat in winter and maize in pre-monsoon. In these districts also commercial vegetable cultivation was done in many areas. Five irrigation systems from Chitwan, three from Makwanpur and four from Nawalparasi constituted the sample. Some of the systems included were agency initiated though they are at present managed jointly by users. Drought and occasional flooding was common.

In **Thailand** we focused on 13 districts of Chiang Mai province in Northern Thailand, covering Ping and Kok River Basins. Ten districts were included from 6 tributaries in the Ping River basin, namely, Mae Wang, San Pa Tong and Saraphi disrricts from Mae Khan sub-basin, Chom Thong district from Mae Klang sub-basin, Mae Tang and Sansai districts from Mae Tngat sub-basin, Doi Saket and San Kampaeng districts from Mae Kuang sub-basin, Mae Rim district from Mae Tang sub-basin, and Mae Cham district from Mae Rak sub-basin. Likewise, three districts were included from Fang sub-basin of Kok River Basin, namely, Fang, Mae Ai and Chaipraklan districts. Rice was the main crop grown in the study districts irrigated mainly by small-scale irrigation systems, but some large systems are also present in the area. The study districts represent both hills and plains physiographic zones, which face frequent flash floodings and extended dry spells. From the study, the major source of water used for all 30 irrigation systems is surface water. This referred to both the head works of the systems and other sources from which water is made available for irrigation. When surface water is the primary source, most type of surface water is river. There were 2 irrigation systems that had the dam for other storage facilities at irrigation system level that can control the flow of units in the production, distribution and appropriation resources for the benefit of all users.

In **Colombia** we sampled three of the main regions of rice production in the country according to the National Federation of Rice Producers (FEDEARROZ). We chose these regions in order to include a representative sample of each different irrigation methodologies: irrigated vs secano (non-irrigated); and for different levels of technology: manual vs mechanized. Irrigated rice production refers to the use of an irrigation system structure to irrigate crops independent on the weather conditions, while secano rice production refers to the use of rain and flood seasons to irrigate crops. Manual rice production refers to the production of rice without using machines while mechanized rice production refers to the opposite case.

The department of Sucre is part of the Costa Norte region were rice production is characterized by small farmers that grow rice manually by secano (dry farming), this area represents only 16% of Colombia's sown area though it holds more than 60% of rice farmers. The departments of Meta and Casanare are part of the Llanos region were rice production is characterized by large farmers that grow mechanized rice by secano, this area represents 35% of Colombia's sown area though it holds less than 10% of rice farmers. Because rice production in these two areas is grown dependent on rain and flood seasons farmers only grow rice during one semester of the year. The departments of Huila and Tolima are part of the Centro region were rice production is

characterized by medium size farmers that grow mechanized rice with irrigation, this area represents 36% of Colombia's sown area and holds around 26% of rice farmers. Inside this region eight different irrigation system associations are established mostly around rice growing, we worked with seven of them. Because irrigation in these region is more independent on weather conditions rice is commonly grown uninterruptedly during the entire year.
To choose subjects for our experiments, in the departments of Sucre, Meta and Casanare we worked in different municipalities and run one or two sessions per municipality. We run six sessions in three municipalities for Sucre, three sessions in three municipalities for Meta and six sessions in six municipalities for Casanare. In the departments of Huila and Tolima we worked with the irrigation system associations which in some cases covered different municipalities hence we didn't use this distinction. We run twelve sessions in total and the number of sessions run on each irrigation system depended on the size of the irrigation system, therefore we have small irrigation systems such as Asoporvenir with 130 families were only one session was run, and other large irrigation systems such as Usocoello with 1800 users were we run four sessions.

In **China**, we chose two provinces located in the Yangtze River basin, the most important rice production area in China. The two provinces are Sichuan province, located at the upper reach of Yangtze, and Jiangxi province, located at the lower reach of Yangtze. Sichuan has the largest irrigation district and the longest irrigation history in China, Dujiangyan, an irrigation infrastructure built in 256 BC. The landscape in Sichuan is dominated by mountainous region combined with plain area. It is ranked in the 6th place in China's rice production. Jiangxi is one of major irrigation districts in China. Its irrigation systems are characterized by large irrigation systems combined with small-scale mountainous irrigation systems. It is ranked in the 3rd place in China's rice production.

To choose subjects for our experiments, in each province, we choose three counties; in each county, we choose 2-3 townships. The main consideration for choosing the townships was the irrigation system types, i.e. mountainous versus plain irrigation systems. In the end, in each province, 15 communities were chosen and therefore we had 30 communities in total from six counties in two provinces. In each community, we used community rosters to choose 16-20 families engaged in agricultural production. The family heads of selected families were asked to participate in our experiments. They were real irrigation water users.

Statistics of communities at the country level

In table S2 below we show the demographics for our total sample of 2,147 people, broken down by country in our 118 communities. About 20% of them were female, with an average age of 50 years. Most participants live in a household of about 4 members, and about 50% of the participants have no more than primary education, although this differs from 40% in Nepal to 70% in Thailand. More than 80% of the participants have lived their whole life in the community.

The farm sizes on average are 1 ha in the Asian countries while being about 26 hectares in Colombia. In Colombia farmers are more likely to rent land compared to the Asian countries where most were land owners.

	Colombia	**China**	**Nepal**	**Thailand**
Session {N sample}	28 {459}	30 {580}	30 {519}	30 {589}
Group size {min/max}	16.39 {12/23}	19.33 {16/20}	17.30 {12/23}	19.63 {15/24}
Age (std.dev)	47.28 (13.32)	54.40 (11.91)	44.84 (15.25)	55.90 (10.90)
Sex (Female = 1)	16.85%	22.42%	20.16%	20.88%
Education	3.35 (1.57)	3.25 (1.16)	3.14 (1.57)	2.85 (1.22)
Land (ha)	33.29 (101.29)	0.92 (3.12)	0.68 (0.69)	1.12 (1.46)
Household size	4.17 (1.83)	3.86 (1.99)	5.87 (3.79)	4.00 (1.46)
Year Income USD	15849.2 (32008)	7187.3 (11136)	3549.7 (4793.5)	4402.4 (4646.5)
Trust Index	0.65 (0.15)	0.70 (0.15)	0.66 (0.13)	0.62 (0.14)
Risk index	0.914 (0.815)	1.180 (0.987)	0.959 (0.832)	0.934 (0.792)
Prediction (Baseline round)	60.07% (22.14%)	70.67% (18.89%)	65.83% (19.64%)	65.33% (18.26%)
Community size (hhs)	13981.5 (29636.1)	545.4 (198.2)	961.27 (2280.2)	196.97 (104.8)
Income from outside %	24.3 (15.2)	42.5 (20.8)	21.77 (18.67)	15.83 (14.98)
DESCRIPTIVE STATISTICS FOR EXPERIMENTAL DECISIONS				
Prediction (Private Risk)	60.11% (20.54%)	68.91% (18.33%)	62.46% (20.06%)	64.98% (19.00%)
Prediction (Collective Risk)	57.24% (21.88%)	66.55% (20.24%)	60.58% (20.80%)	61.81% (19.27%)

Table S2: Descriptive statistics for the four countries. For each variable we provide the average value, the standard deviation between () brackets.

Perceived Threats

Small-scale irrigation communities experience many challenges as discussed in the introduction of the main article. What is the perception of the participants in the experiments? As part of the survey we asked them to rate the level of threat to their livelihood on a 5 point scale for a number of topics (Table S3). Most of the topics are perceived as more than medium threat. For Colombia the volatility of prices are the main threat, while for China the main threats are decay of the irrigation infrastructure and natural disasters. For Nepal the main threats are volatility of prices and the decay of the irrigation infrastructure, while for Thailand the main threats are the volatility of prices.

	Colombia	**China**	**Nepal**	**Thailand**
Changes in weather patterns	4.462 (0.832)	3.477 (1.245)	3.467 (1.033)	3.549 (1.023)
Changes in governmental regulations and subsidies	3.212 (1.680)	2.880 (1.345)	3.294 (0.983)	3.218 (1.079)
Volatility of input prices	4.519 (0.747)	3.550 (1.173)	4.075 (1.051)	4.038 (0.880)
Volatility of market prices products	4.680 (0.686)	3.556 (1.222)	4.192 (0.746)	3.909 (0.981)
Outmigration	3.220 (1.482)	2.649 (1.438)	3.731 (1.019)	2.902 (1.040)
Decay of irrigation infrastructure	4.256 (1.225)	3.815 (1.350)	3.983 (1.284)	3.406 (1.215)
Natural disasters causing floodings	3.567 (1.611)	3.834 (1.403)	3.613 (1.519)	3.045 (1.315)

Table S3: The mean response to the survey question to rate the level of threat to their livelihood as (1) very low, (2) low, (3) medium), (4) high, (5) very high. Between () we list the standard deviation.

IV. THEORETICAL PREDICTIONS

The public goods game used in this study is a variation of the "Voluntary Contribution Mechanism" (8), but simplified given the levels of literacy and logistics in the field where hand-run experiments are required. Our baseline round is one of a pure public goods problem. The so called **Voluntary Contributions Mechanism (VCM)** has been widely used in experimental economics (9) to study the problem of collective action under a simple linear production function of the public good, proportional to the individuals' voluntary contributions. For the particular case here, and knowing the challenges regarding literacy and simplicity, we have used a simple dichotomous cooperation decision-making setting (10) that requires only use of pencil & paper, in the following way:

In our design each player *i* of *m* players has a choice set of two options, $x_i = \{0,1\}$ keep or contribute a token to a public fund or project. If the token is kept it yields a payoff *p* to player *i* only. If the token is invested in the group project, it yields a payoff of *a* to every player *j* including *i.* Summarizing, the payoffs function is given by: $y_i = p(1 - x_i) + a\left(\sum_{j=1}^{m} x_j\right)$. If we analyze the ratio of the marginal return from the private account to the marginal return from the public account we obtain $\frac{\left(\partial y_i \Big/ \partial\left(\sum_{j=1}^{m} x_j\right)\right)}{-(\partial y_i / \partial x_i)} = \frac{a}{p}$. This is the MPCR (Marginal per Capita Return of the public account to the private account) as defined by (11). As long as the MPCR < 1, there will be no incentive to contribute to the group account and therefore the Nash strategy will be $x_i^{nash} = 0$, resulting in a socially inefficient outcome. In such cases, each player obtains $y_i = p$, and the group outcome would be $\sum y_i = mp$. Basically each token in the group account implies a foregone income of (*p-a)* given that no contract has been written between *i* and the rest of the players. However if every player were to contribute to the group account, $x_i^{soc.opt} = 1, \quad \forall i = 1,,m$, the social optimum is obtained. In this case the earnings for each player are $y_i = ma$, and the group outcome would be $\sum y_i = m^2 p > mp$.

In our particular design, we propose to recruit a number of approximately *m=20* players[1] and assign values for a=1 and p=8 obtaining a MPCR of 0.125. Notice, as long as we have more than 8 players in a session, the problem remains a cooperation problem where the Nash dominant strategy is not to contribute to the public fund. If exactly 8 players were to contribute, a player willing to cooperate would be indifferent between keeping the token and contributing to the group account, that is, the critical mass for the collective action problem, expressed as a percentage, would be 8/20=40% of the group size.

---

[1] This design allows for a variable number of m, as long as m>8.

If a participant is risk neutral selfish and rational, the expected utility for all three treatments are the same and thus we will expect no difference between the treatments and the baseline case.

The collective action problem studied in this paper involves two kinds of uncertainties: on the one hand, the returns on the private or the group accounts are uncertain (i.e. the return from keeping the token can be, with a 50/50 chance, 0% or 200% of the $8 from the baseline case; likewise, the return on the group account can be 0% or 200% of the $1 from the baseline case). On the other, individuals do not have certainty about the actions of the others in the group as the game is played simultaneously and no communication is allowed among the group members.

Regarding the first level, we can expect with confidence that most players perceive this as a fair coin toss whereas in the second case, their distribution of probabilities for the number of other cooperators is not necessarily uniform and it may change across group members. Instead, each player has a personal set of priors about the expected rate of cooperation for the rest of the group, given the personal good-will accounting each may have of the others given their reputation and history of interactions and their capacity to read others' intentions in these kinds of social interactions (12,13). Thus, each group member in our experiment will have a personal probability distribution of the expected rate of cooperation in her group and such priors will be determined by the level of trust of this player towards the rest of the group and the expected trustworthiness of the others. Since we elicit in each round the expected fraction of players in the group that would invest their token in the group fund, we can have a proxy for such probability distribution, as well as a measure of the general trust towards others in the community, which we elicited in the post-game survey. Furthermore, we have a measure of the preferences for risk for each player to control for the value that each player places on uncertainty thanks to the lottery experiment after the contributions game is completed.

These two sources of uncertainty are the reasons we claim cooperation is so fragile in a more globalized world in which farmers are facing risk about the possible outcomes regarding shocks in their natural environment or the prices for outputs and inputs they face. Experimental evidence shows that stochasticity in the payoffs in these dilemmas can erode the rate of cooperation (14).

Notice, however, that our two types of uncertainties (in the actions of others and in the returns on the investments) interact in asymmetric ways in our three experimental treatments (BaseLine, PrivateRisk and CollectiveRisk). In the baseline treatment we have only the source of uncertainty regarding the contributions of the others. In the treatment where the return on the private investment (keeping the token), these uncertainties are, in a way, additive. But in the case of uncertainty in the return from the group investment and the uncertainty on the actions of others, such effect is multiplicative, enhancing the negative impact that such uncertainty may have in the willingness to cooperate by each player. Such multiplicative effect does not exist in the case of a risky private investment.

$$(Baseline)Y_i = p \cdot (1 - x_i) + a \cdot \left(\sum x_j\right), donde\ p = \$8, a = \$1$$

$$(PrivateRisk)Y_i = 0.5 \cdot [0 \cdot (1 - x_i)] + 0.5[16 \cdot (1 - x_i)] + 1 \cdot \left(\sum x_j\right)$$

$$(CollectiveRisk)Y_i = p \cdot (1 - x_i) + 0.5 \cdot \left[0 \cdot \left(\sum x_j\right)\right] + 0.5 \cdot \left[2\left(\sum x_j\right)\right]$$

Therefore, we should expect that the level of cooperation should be lower for the case where the uncertainty is placed in the group fund as opposed to the risk in the private investment of the token.

V. ROBUSTNESS CHECKS

In Table S4 we show the mean and standard deviation for the fraction of investment in the four countries for each treatment.

	Fraction of cooperators		
	Round 1	Private Risk	Collective Risk
Colombia	0.522 (0.181)	0.539 (0.145)	0.429 (0.161)
China	0.759 (0.139)	0.745 (0.138)	0.675 (0.139)
Nepal	0.791 (0.123)	0.706 (0.140)	0.645 (0.145)
Thailand	0.662 (0.134)	0.712 (0.126)	0.598 (0.134)

Table S4. Fraction of cooperators by round, by country with the standard deviation of the community level data between brackets.

Using the Wilcoxon signed-rank test we test whether communities have a lower probability investing in the rounds with risk. With private risk, we see that only for Nepal there is a significant effect (P<0.01). When there is a risk at the group level we observe a significant reduction of investments for each country.

	|Round 1 – Private Round|	|Round 1 – Group Round|	|Private Round – Group Round|
Colombia	Z=-0.548 P=0.5840	Z=1.914 P=0.0557	Z=2.563 P=0.0104
China	Z=1.072 P=0.2837	Z=2.729 P<0.01	Z=-2.100 P=0.0357
Nepal	Z=-1.884 P=0.0596	Z=2.593 P<0.01	Z=3.777 P<0.01
Thailand	Z=3.002 P<0.01	Z=4.095 P<0.01	Z=1.925 P=0.0542

Table S5: Wilcoxon signed-rank test results on differences between treatments at community level

Expectations

The prediction on the level on investment of the group is highly correlated with the decisions of the participants. What determine the expectations? In Table S6 we show two models to explain the level of expected investments in round 1. We see that group size matters, namely bigger groups lead to higher percentage of investments. In bigger communities, participants have a slightly lower level expectation. Furthermore, we find that the percentage of the community who earn wage outside the community has a positive relationship with expected investment rate. We also see that more self-sufficient communities have a higher expectation of the investments by the group. With risk, participants expect a lower level of investments, especially with collective risk. When we include interaction effects of treatment and integration with the market, we don't find significant interaction effects.

	Model 1	Model 2
Constant	2.0143 (0.0907)***	1.9921 (0.1067)***
Group size	0.0195 (0.0428)***	0.0195 (0.0043)***
Community size (1000 hhs)	-0.0015 (0.0009)*	-0.0015 (0.0009)*
Outside %	0.0014 (0.0006)**	0.0023 (0.0010)**
Crops for market	-0.0000 (0.0004)	-0.0003 (0.0006)
Food from community	0.0038 (0.0005)***	0.0041 (0.0008) ***
Private Risk	-0.0534 (0.0284)*	0.0136 (0.1016)
Collective Risk	-0.1983 (0.0284)***	-0.1987 (0.1016) **
Outside % * Private Risk		-0.0016 (0.0014)
Outside % * Collective Risk		-0.0016 (0.0014)
Crops for market * Private Risk		0.0008 (0.0009)
Crops for market * Collective Risk		0.0002 (0.0009)
Food from community * Private Risk		-0.0014 (0.0014)
Food from community * Collective Risk		-0.0016 (0.0014)
Number of observations	6364	6364
Wald Chi2	187.49 (p<0.01)	193.02 (p<0.01)
-Log likelihood	-8529.7258	-8527.041
Akaike Information Criterion	17075.45	17082.08
Baysian Information Criterion	17129.52	17176.7
Var (session)	0	0
Chi2	0 (p=1)	0 (p=1)

Table S6: Results from a multilevel logit regression on individual level data explaining the expected investment by the group to invest in the public good.

Expectations and collective risk

Regarding the sequence of rounds for Baseline-CollectiveRisk-PrivateRisk vs Baseline-PrivateRisk-CollectiveRisk, our data analysis does not show an order effect of relevance for the interpretation of the results. In both rounds, 2 and 3, we find that the fraction of individuals investing in the group fund is lower when the risk involves the public good, that is, the group fund (p-value< 0.01 (Mann-Whitney test), as it will be shown below in the supporting regression analysis and in the main text (See Figure S8).

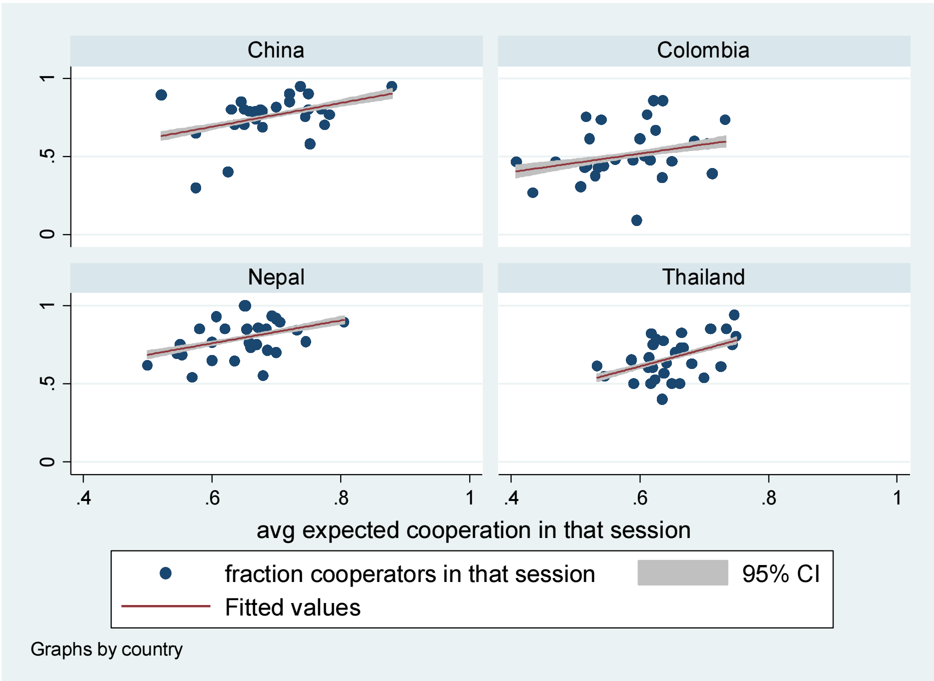

Figure S8. Actual fraction of cooperators and expected fraction of cooperators for baseline round. Pearson correlation coefficients: China=0.4508 (p-value=0.0000); Colombia 0.2332 (p-value=0.0328); Nepal=0.4025 (p-value=0.0001); Thailand=0.2120 (p-value=0.0487).

As discussed before, the lack of information about the decisions of the others in each round brings in a level of uncertainty. However, each player will have a prior about the probability distribution in the group of peers, since they know each other in their respective communities. Such prior will affect, as shown in the results before, the willingness to cooperate and thus the aggregate level at the community level. However, we find that the risk involved in the uncertain returns on the public fund (CollectiveRisk) also affects the predictions or expectations about the others in the group harming in an indirect manner the engine of reciprocal expectations about the cooperation of the others. Figure S9 shows the average levels of expectations for the group and private risk options and once again in all cases the expected cooperation of others decreases with the scenario of a risky public fund.

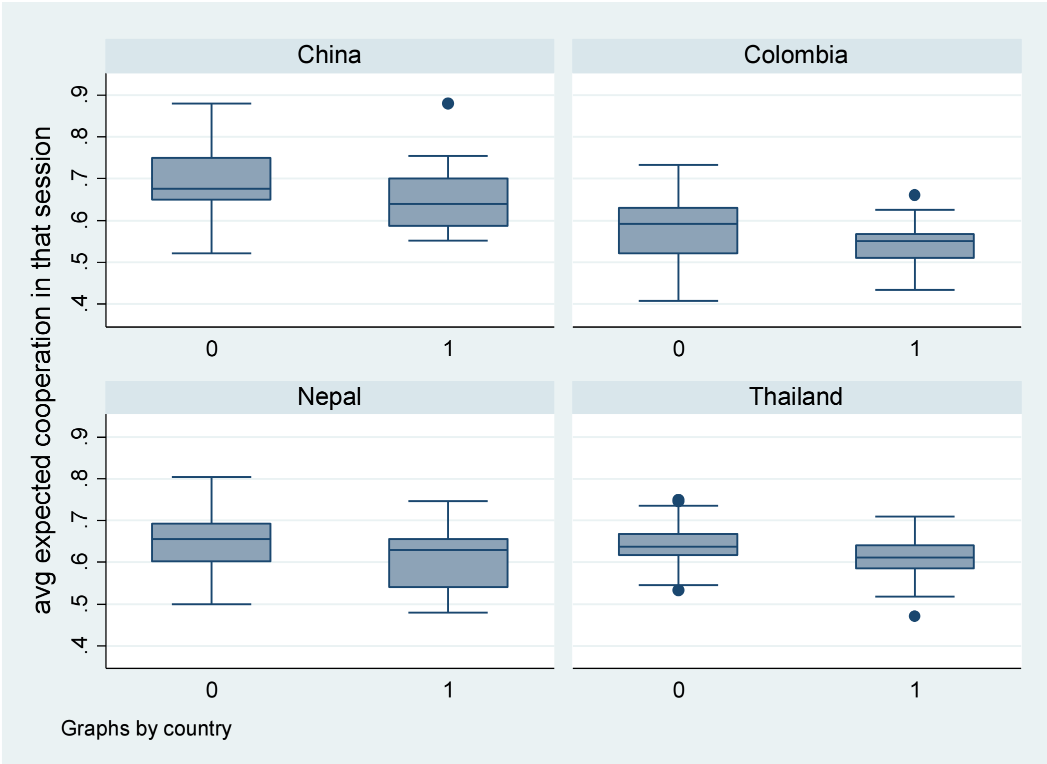

Figure S9. Average expected cooperation for the BaseLine (0) round and the CollectiveRisk (1) treatments.

As a counterfactual check, we also ran a test of differences between the baseline round and the private risk round, for the expected levels of cooperation. Figure S10 shows the means and dispersions. All non-parametric tests confirm no statistical differences.

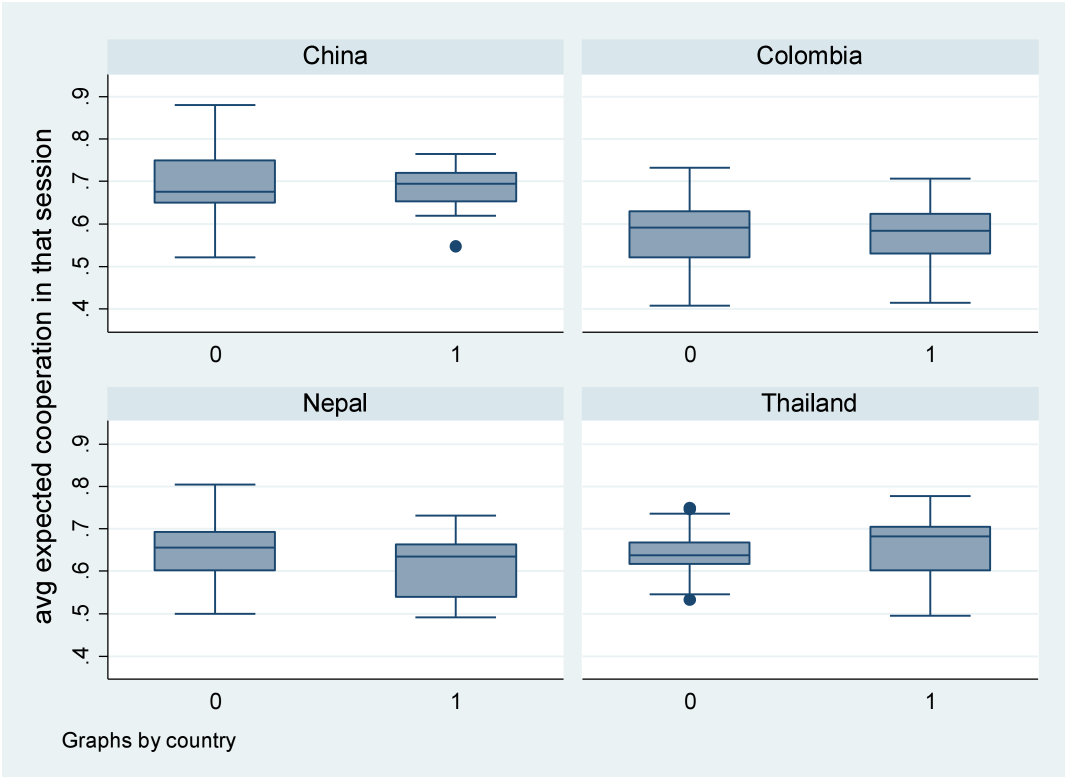

Figure S10. Average expected cooperation for the BaseLine (0) round and the PrivateRisk (1) treatments.

We present now in Table S7 a series of results for the multilevel regression analysis at the individual level. The dependent variable is the decision to invest (=1) or not invest (=0). Because of the binary nature of the decision we use a logit regression. We combine the observations for the 3 rounds, leading to 6441 observations. In Model 1 provides an analysis of the experimental conditions. We find that in larger groups participants are more likely to cooperate, and we find that the level of investment in the CollectiveRisk treatment is lower.

In Model 2 we add community level metrics. We now see that the CollectiveRisk dummy becomes insignificant which means that the different outcomes for the CollectiveRisk treatment are explained by the community level metrics. The size of the community has no significant impact, but how they interact with the broader economy does. We find that communities with a higher level of the population having outside wage options lead to lower level of investments in rounds with the CollectiveRisk treatment. Furthermore, we find that the "fraction of crops grown for the external market" has a significant negative effect on the likelihood a person invests in the public good. Similarly, the "share of food produced by the own community" has a significant positive effect on the likelihood a person invests in the public good. Both effects suggest that more integration with the market leads to a reduction of investments in the public good. Finally, we find that for the PrivateRisk treatment the variable "fraction of crops grown for the external market" has a positive effect.

Model 3 has one additional variable compared to Model 2, the expectation individuals have on the investment level of the group. This has a strong positive effect on the likelihood individual invest in the public good, but the significant effects found in the previous analyses hold. The results of Model 3 are visualized in the main article of this publication (Figure 4). Models 4 and 5 are additional robustness tests. In Model 4 we include various individual attributes. We find that older individuals and female participants invest less. Risk aversion has a modest positive effect on the likelihood of investments. However, trust, income and education do not have significant effects. In Model 5, we include country dummies, and we see that the Asian countries invest significantly more than Colombia. Nevertheless, the main findings from Model 3 remain to hold up.

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	0.0843 (0.3710)	-0.3935 (0.3808)	-1.1290 (0.3749)***	-1.0203 (0.4867)**	-1.0035 (0.5131)**
Private Risk	-0.0370 (0.0680)	-0.11338 (0.2475)	-0.1386 (0.2503)	-0.1856 (0.3062)	-0.1944 (0.3082)
Collective Risk	-0.4339 (0.0665)***	-0.2992 (0.2419)	-0.2288 (0.2448)	-0.1413 (0.3005)	-0.1513 (0.3022)
Group size	0.0412 (0.0199)**	0.0428 (0.0171)**	0.0343 (0.0165)**	0.0280 (0.0188)	0.0166 (0.0193)
Community size (1000 hhs)		0.0145 (0.0345)	0.0199 (0.0333)	0.0074 (0.0416)	0.0352 (0.0408)
Outside %		0.0024 (0.0032)	0.0015 (0.0031)	0.0020 (0.0035)	-0.0011 (0.0037)
Outside % * Private Risk		-0.0044 (0.0034)	-0.0042 (0.0035)	-0.0025 (0.0039)	-0.0025 (0.0039)
Outside % * Collective Risk		-0.0083 (0.0034) **	-0.0079 (0.0034)**	-0.0073 (0.0038)*	-0.0073 (0.0038)*
Crops for market		-0.0048 (0.0020)**	-0.0048 (0.0019)**	-0.0045 (0.0022)**	-0.0011 (0.0023)
Crops for market * Private Risk		0.0042 (0.0021)**	0.0041 (0.0022)*	0.0041 (0.0025)*	0.0031 (0.0025)*
Crops for market * Collective Risk		0.0014 (0.0021)	0.0014 (0.0021)	0.0002 (0.0024)	0.0002 (0.0024)
Food from community		0.0106 (0.0024)***	0.0093 (0.0023)***	0.0082 (0.0027)***	-0.0003 (0.0036)
Food from community * Private Risk		-0.0010 (0.0026)	-0.0006 (0.0026)	0.0002 (0.0031)	0.0003 (0.0031)
Food from community * Collective Risk		-0.0002 (0.0025)	-0.0004 (0.0026)	0.0002 (0.0031)	0.0003 (0.0031)
Prediction			0.3843 (0.0308)***	0.4570 (0.0372) ***	0.4520 (0.0371)***
Age				-0.0085 (0.0031)***	-0.0076 (0.0031)**
Sex (Female = 1)				-0.1932 (0.0877)**	-0.01782 (0.0876)**

Education				0.0445 (0.0470)	0.0429 (0.0469)
Land (ha)				0.0951 (0.0443)**	0.0909 (0.0442)**
Household size				0.0133 (0.0166)	0.0079 (0.0169)
Income				0.0471 (0.0402)	0.0522 (0.04019)
Trust				0.2851 (0.2522)	0.2951 (0.2538)
Risk aversion				0.1184 (0.0687) *	0.1172 (0.0686)*
Risk aversion * Private Risk				-0.1046 (0.0938)	-0.1049 (0.0937)
Risk aversion * Collective Risk				-0.0813 (0.0921)	-0.0813 (0.0921)
China					0.9169 (0.2706)***
Nepal					0.4647 (0.1835)**
Thailand					0.9089 0.2307)***
Number of observations	6440	6374	6363	4830	4830
Wald Chi2	57.52 (p<0.01)	111.01 (p<0.01)	262.46 (p<0.01)	252.02 (p<0.01)	271.28 (p<0.01)
-Log likelihood	3990.8433	3926.4637	3839.457	2855.9939	2848.5182
Akaike Information Criterion	7991.687	7882.927	7710.914	5763.988	5755.036
Baysian Information Criterion	8025.538	7984.327	7819.046	5932.536	5943.032
Var (session)	0.3131 (0.0534)	0.1812 (0.0357)	0.1592 (0.0329)	0.1679 (0.0385)	0.1319 (0.0335)
Chi2	249.78 (p<0.01)	114.90 (p<0.01)	93.10 (p<0.01)	65.82 (p<0.01	44.95 (p<0.01)

Table S7: Results from a multilevel logit regression on individual level data explaining the decision to invest in the public good.

In the next Table S8 we perform analysis at the individual level for each of the three treatment rounds separately. We find that prediction is the most significant factor in each treatment round. We also find that the percentage of wage labor outside the community is only significant in the CollectiveRisk treatment. Higher percentage of wage labor outside the community reduces the investment in the public good if there is a collective risk. We do find significant effects for the integration with the market for the individual treatment rounds.

	Base round	PrivateRisk	CollectiveRisk
Constant	-0.9096 (0.5063)*	-1.0126 (0.4790)**	-1.7442 (0.4461)
Group size	0.0113 (0.0242)	0.0287 (0.0227)	0.0565 (0.0210)***
Community size (1000 hhs)	0.0347 (0.0497)	-0.0563 (0.0478)	-0.0289 (0.0442)
Outside %	0.0014 (0.0034)	-0.0026 (0.0032)	-0.0065** (0.0029)
Crops for market	-0.0051 (0.0021)**	-0.0009 (0.0020)	-0.0033 (0.0018)*
Food from community	0.0095 (0.0026)***	0.0093 (0.0024)***	0.0082 (0.0022)***
Prediction	0.4703 (0.0560)***	0.3105 (0.0548)***	0.3951 (0.0503)***
Number of observations	2119	2119	2122
Wald Chi2	100.44 (p<0.01)	52.42 (p<0.01)	101.28 (p<0.01)
-Log likelihood	1222.864	1272.674	1354.873
Akaike Information Criterion	2461.727	2563.348	2725.745
Baysian Information Criterion	2507.008	2608.618	2771.026
Var (session)	0.2303 (0.0720)	0.1859 (0.0613)	0.1372 (0.0513)
Chi2	23.37 (p<0.01)	19.79 (p<0.001)	13.48 (p<0.01)

Table S8: Results from a multilevel logit regression on individual level data explaining the decisions to invest in the public good for each of the 3 treatments separately.

Below Table S9 shows a new model version using group level data. The results supports largely the individual level data results. At the group level the prediction of the group level investment, and the market integration indicators are highly significant factors. The interaction effect between the percent of the population having wage labor outside the community and the group level risk is less significant in the group level analysis than the individual level analysis. Finally, if we include dummy variables for countries, the impact of market integration on the share of investments is not significant anymore.

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	0.5241 (0.0567)***	0.4400 (0.0695)***	0.1579 (0.0779)**	0.1602 (0.1520)	0.3128 (0.1614)*
Private Risk	-0.0085 (0.0217)	-0.01647 (0.0698)	-0.0240 (0.0657)	-0.0555 (0.0875)	-0.0548 (0.0849)
Collective Risk	-0.0968 (0.0217)***	-0.0918 (0.0698)	-0.0670 (0.0658)	-0.1132 (0.0876)	-0.1150 (0.0849)
Group size	0.0089 (0.0030)***	0.0093 (0.0028)***	0.0058 (0.0026)**	0.0053* (0.0029)	0.0001 (0.0032)
Community size (1000 hhs)		0.0038 (0.0055)	0.0058 (0.0052)	0.0031 (0.0054)	0.0085 (0.0054)
Outside %		0.0005 (0.0007)	0.0001 (0.0006)	0.0003 (0.0007)	-0.0001 (0.0007)
Outside % * Private Risk		-0.0009 (0.0010)	-0.0007 (0.0009)	-0.0008 (0.0009)	-0.0008 (0.0009)
Outside % * Collective Risk		-0.0018 (0.0010)*	-0.0014 (0.0009)	-0.0015 (0.0009)*	-0.0016 (0.0009)*
Crops for market		-0.0011 (0.0004)***	-0.0011 (0.0004)***	-0.0010 (0.0004)**	-0.0005 (0.0004)
Crops for market * Private Risk		0.0007 (0.0006)	0.0006 (0.0005)	0.0006 (0.0005)	0.0006 (0.0005)
Crops for market * Collective Risk		0.0003 (0.0006)	0.0002 (0.0005)	0.0003 (0.0005)	0.0003 (0.0005)
Food from community		0.0020 (0.0005)***	0.0014 (0.0005)***	0.0013 (0.0005)***	-0.0002 (0.0006)
Food from community * Private Risk		-0.0001 (0.0007)	0.0001 (0.0007)	0.00004 (0.0007)	0.00001 (0.0007)
Food from community * Collective Risk		0.0004 (0.0007)	0.0003 (0.00007)	0.0003 (0.0007)	0.0003 (0.0007)
Prediction			0.1489 (0.0223)***	0.1545 (0.0228)***	0.1321 (0.0226)***
Age				-0.0002 (0.0012)	-0.0010 (0.0014)
Sex (Female = 1)				0.0394 (0.0555)	0.0235 (0.0544)
Education				0.0018 (0.0011)*	0.0015 (0.0010)
Land (ha)				-0.0003 (0.0011)	-0.0017 (0.0011)
Household size				0.0163 (0.0076)**	0.0126 (0.0098)
Income				0.0015 (0.0010)	0.0021 (0.0010)**
Trust				-0.0977 (0.1496)	-0.0273 (0.1591)
Risk aversion				-0.0208 (0.0461)	-0.0492 (0.0453)

Risk aversion * Private Risk				0.0340 (0.0634)	0.0344 (0.0614)
Risk aversion * Collective Risk				0.0512 (0.0635)	0.0491 (0.0615)
China					0.2014 (0.0451)***
Nepal					0.1251 (0.0318)***
Thailand					0.1593 (0.0385)***
Number of observations (groups)	354	351	351	351	351
R^2	0.0866	0.2797	0.3642	0.3906	0.4334
F-test	11.07	10.07	13.75	8.71	9.15
Akaike Information Criterion	-258.9038	-338.0631	-379.8643	-374.7492	-394.3147
Baysian Information Criterion	-243.4266	-284.012	-321.9525	-278.2296	-286.2127

Table S9: Results from a linear regression on community data explaining the fraction of investment in the public good.

We want to show how the change in the reduction of investment is correlated with the level of income from outside the community. We split the data up in three groups (bins) and compare the differences. The sizes of the bins are chosen in such a way that we have about equal number of observations. Hence we can now test the differences between those individuals who are in a community with a small percentage of outside labor, versus those with a high percentage of outside labor. Using the Wilcoxon signed-rank test we test whether the Default round, round 1, is different with the rounds with risk.

	% of outside labor		
	[0,10]	[11,30]	[31,100]
# observations	747	739	661
Default Round	68.41%	66.71%	71.86%
Private Risk	70.01%	66.04%	68.38%
Collective Risk	61.98%	68.38%	56.36%
ΔInvestment Private Risk Test with Default Round	1.61% P=0.3889 Z=0.862	-0.68% P=0.7355 Z=-0.338	-3.48% P=0.0891 Z=-1.700
ΔInvestment Collective Risk Test with Default Round	-6.43% P=0.0011 Z=-3.266	-6.50% P=0.001 Z=-3.281	-15.58% P=0.000 Z=-6.867

Table S10: The different metrics for individuals allocated in 3 similar sized bins using the variable "% of the population involved with outside wage labor".

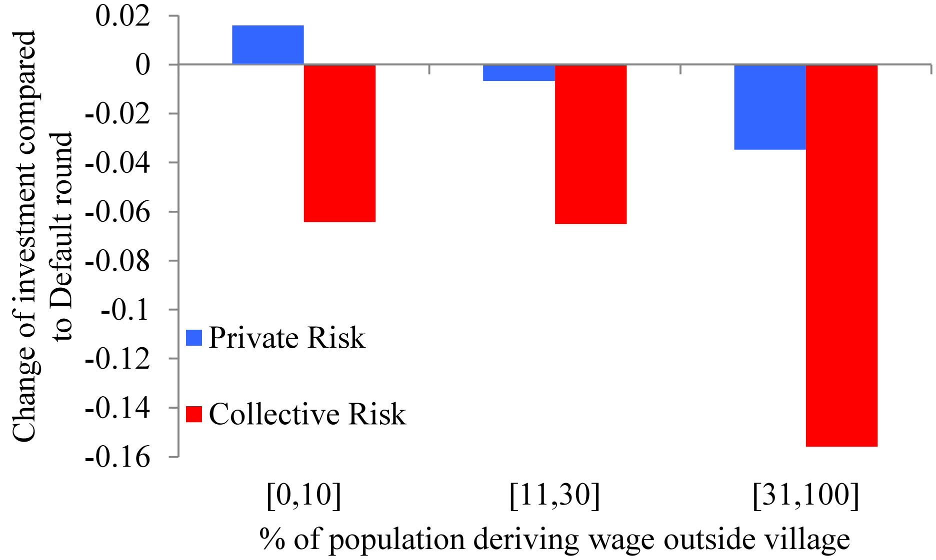

Figure S11. The results from Table S10 graphically explained.

At the community level we test whether the level of reduction in investment in the public good is significantly different between the different bins. Using the Mann-Whitney U-Test (2-tailed), we find that the reduction in investment for the Private Risk rounds are not significantly different (p>0.1). The reduction in investment for the Collective Risk round is also not significantly different between bin [0,10] and bin [11,30] (p=0.9840), but it is significant different between bin [0,10] and bin [31,100] (p=0.0135) and between bin [11,30] and bin [31,100] (p=0.0202). This means that communities with a high percentage of outside labor exhibit significantly lower investment levels in the public good.

	% of outside labor		
	[0,10]	[11,30]	[31,100]
# observations	42	40	36
Default Round	67.73%	66.77%	71.77%
Private Risk	69.75%	65.39%	68.17%
Collective Risk	60.96%	59.63%	55.86%
ΔInvestment Private Risk	2.02%	-1.38%	-3.60%
ΔInvestment Collective Risk	-6.76%	-7.14%	-15.91%

Table S11. The different metrics at the group level when we allocate the groups in 3 similar sized bins using the variable “% of the population involved with outside wage labor”.

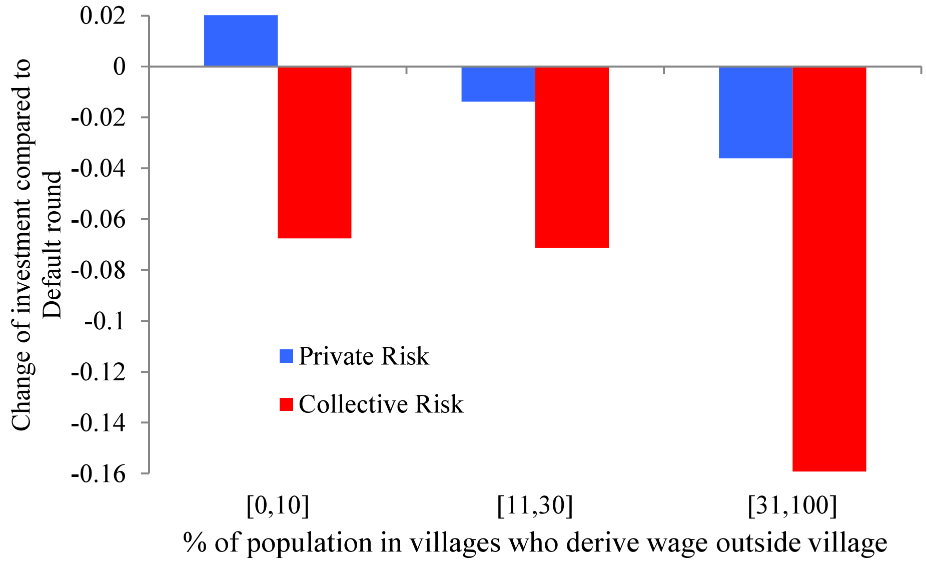

Figure S12. The results from Table S11 graphically explained.

References

1 Binswanger HP (1980) Attitudes towards risk. *American Journal of Agricultural Economics* 62(3): 395-407.

2 Cardenas JC, Carpenter J (2013) Risk attitudes and economic well-being in Latin America. *Journal of Development Economics* 103: 52–61.

3 Holt CA, Laury SK (2002) Risk Aversion and Incentive Effects. *American Economic Review* 92(5): 1644-1655.

4 Harrison GW, Lau MI, Rutstrom EE (2007) Estimating Risk Attitudes in Denmark: A Field Experiment. *Scandinavian Journal of Economics* 109(2): 341-368.

5 Cárdenas JC, Chong A, Ñopoo H (2013) Stated social behavior and revealed actions: Evidence from six Latin American countries, *Journal of Development Economics* 104(C): 16-33.

6 Chellaney, B (2013) *Water, Peace, and War: Confronting the Global Water Crisis.* (Rowman & Littlefield, Lanham, MD).

7 Fant C, Schlosser CA, Gao X, Strzepek K, Reilly J (2016) Projections of Water Stress Based on an Ensemble of Socioeconomic Growth and Climate Change Scenarios: A Case Study in Asia. *PLoS ONE* 11(3): e0150633.

8 Marwell G, Ames RE (1979) Experiments on the Provision of Public Goods. *American Journal of Sociology* 84(6): 1335-1360.

9 Ledyard J. (1995) Public Goods. (in *Handbook of Experimental Economics*, eds. J. Kagel, A. Roth, Princeton University Press, Princeton, USA, 111-184.

10 Cárdenas JC, Jaramillo C (2007) *Cooperation In Large Networks*. Documentos CEDE, Universidad de los Andes-Cede.

11 Isaac RM, Walker JM (1988) Group Size Effects in Public Goods Provision. *Quarterly Journal of Economics* 103(1): 179-199.

12 Coricelli G, McCabe K, Smith V (2000) Theory-of-Mind Mechanism in Personal Exchange (in *Affective minds*, eds., G. Hatano, N. Okada, H. Tanabe, Elsevier, Amsterdam, 249-259).

13 McCabe K, Smith V (2001) Goodwill Accounting and the Process of Exchange. (in *Bounded Rationality*, eds., G. Gigerenzer, R. Selten, MIT Press, Cambridge, MA, 319-340).

14 Bereby-Meyer Y, Roth A (2006) The speed of learning in noisy games. *American Economic Review* 96: 1029-1042.