



Is a PD game still a dilemma for Japanese rural villagers? A field and laboratory comparison of the impact of social group membership on cooperation

Yohei Mitani¹

Received: 13 May 2021 / Revised: 21 June 2021 / Accepted: 15 July 2021 /

Published online: 26 July 2021

© The Author(s) 2021

Abstract

Local norms and shared beliefs in cohesive social groups regulate individual behavior in everyday economic life. I use a door-to-door field experiment where a hundred and twenty villagers recruited from twenty-three communities in a Japanese rural mountainous village play a simultaneous prisoner's dilemma game. To examine whether a set of experiences shared through interactions among community members affect experimental behavior, I compare villagers' behavior under in-community and out-community random matching protocols. I also report a counterpart laboratory experiment with seventy-two university student subjects to address the external validity of laboratory experiments. The findings are three-fold. First, almost full cooperation is achieved when villagers play a prisoner's dilemma game with their anonymous community members. Second, cooperation is significantly higher within the in-group compared to the out-group treatment in both the laboratory and field experiments. Third, although a significant treatment effect of social group membership is preserved, a big difference in the average cooperation rates is observed between the laboratory and field.

Keywords Lab-in-the-field experiment · Cooperation · Group identity · External validity · Local norms · Prisoner's dilemma

JEL Codes C91 · C93 · C72

✉ Yohei Mitani
yomitani@gmail.com

¹ Division of Natural Resource Economics, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

1 Introduction

Social interactions and social norms play a central role in decision-making in favor of cooperation and coordination. Lifelong, frequent interactions with community members remind that full cooperation can be supported in the equilibrium of infinitely repeated prisoner's dilemma games. However, there is a lack of experimental evidence to support full cooperation in a well-controlled experimental game. Can local norms of cooperation be replicated in an experiment if subjects bring their shared experience and beliefs into the experiment? Understanding the connection between experimental behavior and the elements of everyday economic life still remains an open question (Henrich et al., 2001). In this paper, I report the first experimental evidence of full cooperation among Japanese rural villagers in a prisoner's dilemma game.

The rural areas of Japan have a number of small and cohesive closed communities, in which neighboring villagers share behavioral norms and expectations over generations. Villagers repeatedly interact over time, coordinate almost every day, and take collective action each season. Most local communities seem to have institutional features that meet the requirements of a possible full cooperation outcome that can be supported in an infinitely repeated prisoner's dilemma game. A growing experimental literature has shown that cooperation only emerges with experience even when games are consistent with equilibrium, risk dominant, and sufficiently incentive compatible (Dal Bó & Fréchet, 2011, 2018). Villagers play super games with their community members since their birth, share the lifelong history of those games, and expect certain interactions for the rest of their life. Combined with experimental evidence summarized by Dal Bó and Fréchet (2018), these features lead us to expect that high cooperation can be observed in those communities. Can pre-existing, shared experiences facilitate coordination on the efficient equilibrium in a prisoner's dilemma experiment? How high can we expect cooperation rates to be? I conduct a door-to-door field experiment in a Japanese rural village to capture a moment of everyday interactions among villagers as a stage game of super games.

Group identity can affect individual behavior toward cooperation and coordination (Efferson et al., 2008; Solow & Kirkwood, 2002). Using minimal groups induced in a laboratory, a series of experimental studies show that subjects behave more pro-social with in-group members and can use a salient group identity to coordinate with in-group members on the better equilibrium in a coordination game (Chen & Chen, 2011; Chen & Li, 2009). In real social groups, their members share a set of beliefs and behavioral norms formed through mutually experienced interactions. Social group identity gives more than just a label but shared expectations based on the history of group members. Bernhard et al. (2006) use several native tribes in Papua New Guinea as social groups and show that in-group favoritism regulates altruistic norm enforcement. Goette et al., (2006, 2012) utilize a randomly assigned social group and find that subjects are more likely to cooperate when they are paired with in-group members. In this paper, I explore how strongly local norms and shared beliefs from everyday interactions

among neighboring villagers affect experimental behavior. I do this by comparing the cooperative behavior in a prisoner's dilemma game when they play with their community members and when they play with out-community members.

A growing body of literature addresses the external validity of laboratory experiments with university student subjects (Snowberg & Yariv, 2021). The question of whether the results from a university-student sample can be generalized to other populations remains a fundamental methodological issue in the field of experimental economics. To test the external validity of the laboratory findings, a series of experimental studies compare the behavior of standard subjects with other subject pools. Some of these studies simply compare experimental measures like cooperation rates between different subject pools (Anderson et al., 2013; Bigoni et al., 2013; Burks et al., 2009). Additionally, few compare the difference of experimental treatment effects, for example, of having an opportunity of punishment between different subject pools (Alm et al., 2015; Bortolotti et al., 2015). For example, the literature suggests that university students are less prosocial compared to non-student subjects (Anderson et al., 2013; Bortolotti et al., 2015; Burks et al., 2009). I report a counterpart laboratory experiment with university student participants, which share the same procedures as the field experiment with villagers. I test the external validity of the experimental treatment effect as well as cooperation rates in both experiments.

I conduct a door-to-door field experiment where 120 villagers were selected from 23 communities in a Japanese rural village to play a simultaneous prisoner's dilemma game. To examine whether experiences shared through lifelong, everyday interactions among community members affect experimental behavior, I compare villagers' behavior under in-community and out-community random matching protocols. I also report a counterpart laboratory experiment with 72 university student subjects to address the external validity of laboratory experiments, or the question of whether the results from a university student sample can be generalized to other populations. The findings are three-fold. First, almost full cooperation is achieved when villagers play a prisoner's dilemma game with their anonymous community members. Second, cooperation is significantly higher in the in-community compared to the out-community treatment in the field experiment with villagers, even though a very high cooperation rate is still observed in the out-community treatment. Third, regarding the external validity, a significant treatment effect of social group membership is preserved in both laboratory and field experiment. Although, a big difference in average cooperation rates is observed.

2 Institutional background: local communities in mountainous villages

The rural areas of Japan have a number of small and tight-closed communities. One of my interests is to explore how local norms of cooperation and experience shared by community members affect behavior in an experimental game.

The study site, Kumakogen town, is located in the center of Ehime prefecture (33° 39' N, 132° 54' E), approximately 600 km southwest of Tokyo (Mitani & Shimada, 2021). The town is very mountainous and has a total area of 583.7 km²,

which is almost ten times bigger than the land area of Manhattan. In 2014, the resident population of the town was 9327 with 45.3% of them older than 65 years of age, indicating that the town faces an aging and shrinking population. In total there are 219 communities in the town, and many communities are remote and isolated.

A mail survey was conducted by approaching all 219 community leaders, and 115 responses were collected with an overall response rate of 52.5%.¹ The survey reveals some institutional background. The median community size is 14 households, which can be small enough for community enforcement to support cooperation as a possible equilibrium outcome in an infinitely repeated prisoner's dilemma game. Even though previous experimental studies show that cooperation will not emerge under anonymous random matching with identified past behavior unless the group size is very small (Camera et al., 2012; Duffy & Ochs, 2009), I consider the community size small enough for cooperation to evolve given the lifelong, frequent interactions within community members.

The survey reveals information about community organizations and their collective activities. 93% of households are members of local community organizations. 95% of organizations have collective activity agreements, and more than 30% of them have an enforcement instrument using a monetary penalty. 41% of organizations own a community forest. Regarding collective decision rules, about 40% of communities employ majority voting, about 35% of them employ consensus decision-making (which requires unanimity to reach a decision), and the rest of them require a majority approval for a leader's proposal. This suggests frequent interactions among community members and sustaining local norms of cooperation.

The survey provides evidence that indicates cohesive interactions among community members over generations. The majority of communities do not have any single immigrated household in their current generation. Only 7% of all households have immigrated into the town in their generation. Thus, most villagers have been members of their community since their birth. Only 3% of them commute outside of the town to work. Most of the villagers are retired or work in agricultural, forestry, or public sector jobs in their residential region. Many communities are remote and isolated in mountain valleys. For example, it takes about 2 h to drive up to some communities from the center of the town, with some communities located an hour away from a state road.

Institutional characteristics of these communities are summarized as follows. First, the community size is small. The size in the sample is slightly bigger than that of the Swiss army training platoons used as social groups in Goette et al., (2012) and much smaller than that of the Israeli kibbutz communities used in Ruffle and Sosis (2006).² Second, local community membership is not self-selected, and most villagers have been members since birth. Unlike in the case of the Israeli kibbutz, where prosocial people usually choose to join the community (Ruffle & Sosis, 2006), self-selection of community membership will not account for prosocial behavior in this

¹ See Mitani et al. (2015) for detail.

² The average community size is 21.8 households in our sample while between 500 and 700 individuals in Ruffle and Sosis' sample and 21 individuals in Goette's sample.

study. Third, villagers are relatively homogeneous among communities in terms of individual, demographic, historical, and cultural characteristics. In addition, there is no indication of competition between communities, like there is among tribes or college fraternities (Bernhard et al., 2006; Kollock, 1998).

I use 120 villagers from 23 communities as the subject pool to explore whether local community membership impacts experimental behavior. This is done by comparing their behavior between in-community and out-community anonymous random matching. Local communities in this study area are naturally formed social groups. The institutional characteristics indicate that community members have mutually established behavioral beliefs and norms of cooperation based on their life-long experience of social interactions and personal history. I assume that most villagers will be certain about the norms and shared beliefs of cooperative behavior when their own community is concerned, but they might be unsure when they interact with villagers outside their community. Even though most communities might have similar behavioral norms of cooperation and most villagers know it, the extent to how obvious villagers mutually share the beliefs regarding other villagers' strategies in an experimental game would vary between when they are paired with in-community members and when paired with out-community members.

Following the community leader survey, we contact the community leaders who revealed their willingness to participate in a door-to-door survey and collect a potential list of household volunteers. We recruit 120 participants from 23 communities based on the available number of households in each community and their available date and time. They are then assigned either in the in-community or out-community treatment.

In the in-community treatment, villagers are matched anonymously but informed that the other players in an experimental group are members of their community. The out-community treatment was the same except that villagers are informed that the others are members of different communities. Because anonymity is preserved, villagers only know whether they are paired with someone from their community or different communities, but not with whom. Despite anonymous random matching, villagers in the in-community treatment may not face strategic uncertainty on the others' choices due to their strong enough shared behavioral norms. While, in the out-community treatment, given the town population size of about 10,000, it might be hard for villagers to believe that the others share the same behavioral norms in the experiment. Note that all communities in the town share relatively homogeneous demographic, cultural, and historical characteristics.

Goette et al., (2006, 2012) argue that membership of naturally formed social groups can be confounded with subjects' individual, demographic, and cultural characteristics. Community membership is not randomly assigned, however, nor does it involve a selection based on individual characteristics. There is little evidence of a difference in demographic and cultural backgrounds among local communities in the town. There should also be a small level of statistical difference in beliefs and norms that are brought into the experiment between treatments because all subjects are recruited from the same pool of villagers. However, information available for a subject regarding the opponents can vary across treatments. In the in-community treatment, subjects know the anonymous opponents should be three

Table 1 Payoff matrix for player i

	$(i - 1)$'s decision	
	Send	Keep
i 's decision		
Send	2000	0
Keep	3000	1000

of their community members so that it can be easy for them to form beliefs of others' choice. On the other hand, in the out-community treatment, subjects know their opponents are not their community members so that they might wonder what norms the opponents might have. The observed difference in cooperation in a prisoner's dilemma experiment between treatments can be attributed to the difference of the shared beliefs and experiences between treatments. I try to see whether a prisoner's dilemma game is still a dilemma for community members.

3 Experimental design

3.1 Game

To test whether a set of experiences shared by neighboring villagers regulate pro-social and indirect trust behaviors, I develop a simultaneous prisoner's dilemma game with four local player interactions around a circle. Four anonymous villagers are arranged in a circle with every villager having two directly connecting villagers to their left and right. These direct villagers are referred to as the posterior (left) and anterior (right) participants in the instructions. A villager i has another villager $(i - 1)$ as her right participant and villager $(i + 1)$ as her left participant (for $1 \leq i \leq 4$, let 0 be 4 and 5 be 1). Each villager participates in a one-shot game, where they are asked to choose simultaneously whether to send their whole endowment of 1000JPY (about 9.5USD at the time of the experiment) to the participant on the left or to keep it in their pocket. The amount sent is doubled. The payoff is determined by own and right villager's choice: $\pi_i = 1000(1 - d_i) + 2000d_{i-1}$ where $d_i = 1$ if the money is sent and 0 otherwise. The payoff matrix for player i is illustrated in Table 1.

Keeping means defection while sending means cooperation.³ The payoffs share the same features as standard two-player prisoner's dilemma in the sense that defection strictly dominates cooperation for each player. This game has a unique Nash equilibrium in which all players defect. Whereas the socially optimal outcome is that all players cooperate. If this is the stage game of an infinitely repeated game,

³ In the instructions, I used "Send" and "Keep" representing cooperation and defection, respectively. Field participants are known to find the instructions and protocol difficult (Cardenas and Carpenter, 2008). I carefully chose wording that minimizes framing effects but also helps villagers' understanding of the game structure. In addition, I found low cooperation rates among university students with the same instructions.

cooperation will be consistent with equilibrium and risk dominant for the probability of continuation greater than $2/3$ in the super game. Pro-social and indirect trust behaviors can be observed by the game in the sense that the participants play the dictator role and simultaneously serve as recipients through an indirect participant.

3.2 Treatments

I use a between-subjects design with two treatments: in-community and out-community. In the in-community treatment, villagers are matched anonymously but informed that the other three players in her experimental group are members of the same community. The out-community treatment is the same except that villagers are informed that the other three players are members of different communities but residents in the town. Despite anonymous random matching, villagers in the in-community treatment may not face strategic uncertainty on the others' choices because of strong enough shared behavioral norms. If this is the case, participants know that an implicitly agreed strategy would be either cooperation or defection based on the history of the super game (i.e. daily economic interactions) outside of the experiment. While, in the out-community treatment, it might be hard for villagers to believe that the others share the same behavioral norms in the strategic environment induced by the experiment.

3.3 Procedures

The door-to-door field experiment was conducted with paper and pencil at the participant's residence. A reminder of an interview date and time was mailed to 120 participants recruited from 23 communities, but they were not informed about the economic experiment in advance. 72 of them were randomly assigned to the in-community treatment (18 groups) while 48 were assigned to the out-community treatment (12 groups). The experiment was carried out by a trained undergraduate experimenter visiting each participant at their home.⁴ 12 sessions were conducted in 2 days with 101 participants in total.⁵ At each session, a maximum of 14 experimenters visited their assigned participant and started the experiment at the same time at different places.⁶ This procedure prevents villagers from communicating with other villagers. The experiment lasted 30–45 min and the subjects earned on average 1901 JPY from this specific experimental game. The economic experiment was followed by a post-experiment interview. Special care was taken to ensure anonymity and

⁴ There were 14 experimenters. Among them, eight were female and six were male.

⁵ 19 of the recruited participants (13 in the in-community treatment and 6 in the out-community treatment) cancelled with short notice or did not show up for their appointment. We implemented all experiments as we planned. Indeed, the experimenter was not able to know whether there were any cancellations in their group at the time of their session. We slightly modified a way to calculate the participants' payoffs in case of cancellation in their group.

⁶ The number of participants in a session ranged from 4 to 14 with median of 10.

Table 2 Individual Characteristics

		Villagers				Students
		Out-Com	In-Com	Diff.*	Total	Total
		Mean (SD)	Mean (SD)	<i>p</i> value	Mean (SD)	Mean (SD)
Trust	One can trust people (4 scale)	2.81 (0.67)	2.75 (0.58)	0.87	2.77 (0.61)	2.53 (0.63)
Reward	Willing to reward others (4 scale)	3.76 (0.43)	3.66 (0.48)	0.28	3.70 (0.46)	3.39 (0.52)
Punish	Willing to punish others (4 scale)	2.07 (0.71)	2.12 (0.83)	0.90	2.10 (0.78)	2.69 (0.82)
GFavor	Scale of in-group favor (+ : in-group, 0: no diff)	0.24 (0.53)	0.42 (0.62)	0.17	0.35 (0.59)	0.33 (0.61)
Age	(years)	67.2 (12.5)	67.9 (11.2)	0.90	67.6 (11.7)	20.4 (1.51)
Gender	Male (dummy)	0.67 (0.48)	0.56 (0.50)	0.28	0.60 (0.49)	0.71 (0.46)
<i>N</i>		42	59		101	72

*Diff. shows *p* values from the Mann–Whitney *U* tests for the comparison of out-community and in-community groups

minimize the experimenter effect. Experimental earnings were mailed to the home address a week after the experiment.

A counterpart laboratory experiment with student subjects was conducted at Kyoto University with 72 undergraduate students. I used a classmate matching protocol as an in-group treatment. In an out-group treatment, 48 undergraduate students were recruited from the general population at Kyoto University through a standard recruiting process while 24 undergraduate students were recruited from a specific freshman class of about 40 students at a specific department in the in-group treatment. Experimental procedure and instructions were identical to those used in the door-to-door field experiment with villagers, except that all experiments were run on a computer in an experimental laboratory with 12 subjects at the same time.

3.4 Subject pools

I collected data from 101 villagers and 72 students. Table 2 provides summary statistics from the participant characteristics gathered by a post-experiment questionnaire, along with the results of balance tests determining whether the mean of each characteristic is statistically the same between treatments in the field. The results indicate that the randomization was properly performed.

“Trust” is a survey measure of general trust. Specifically, this question “In general, one can trust people” was asked using a 4-scale measure from “disagree fully (1),” “disagree somewhat (2),” “agree somewhat (3),” to “agree fully (4).” Fig. 1 shows the results by villagers (navy bar) and students (orange bar), suggesting that villagers tend to show relatively higher measures of trust than students. The variables “Reward” and “Punish” were derived from the question “You are willing to reward (punish) others in return for kind (unfair) treatment” with a 4-scale measure. Figures 2 and 3 suggest that villagers are more likely to agree with rewarding while they disagree with punishing when compared with students. “GFavor” was

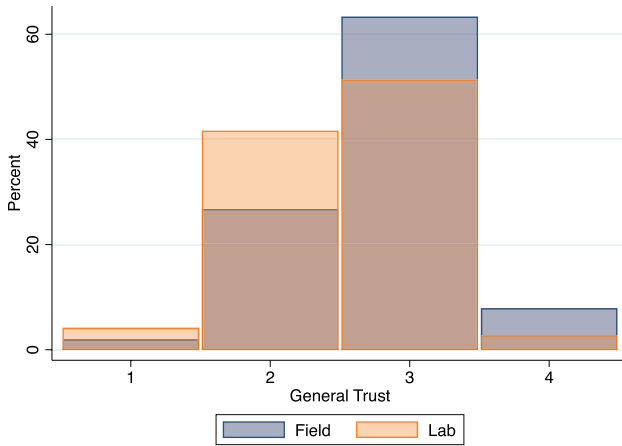


Fig. 1 General Trust

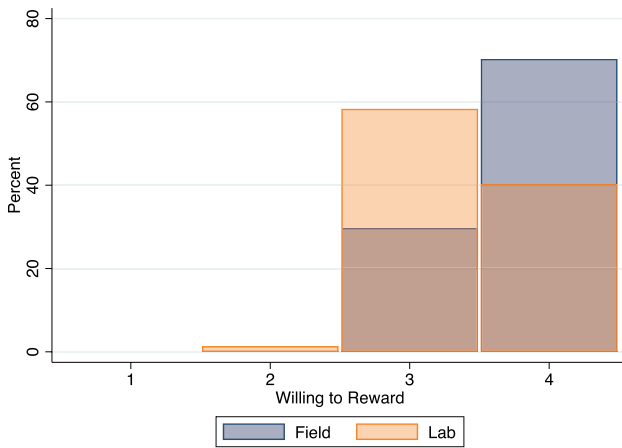


Fig. 2 Willingness to Reward

constructed by combing answers to the questions with a 4-scale measure: “You are willing to help neighbors who in need” and “You are willing to help strangers who in need.” This measure of 0 is consistent with no evidence of in-group favoritism while positive scores are consistent with in-group favoritism. Figure 4 shows no difference between villagers and students.

4 Results

The findings from the door-to-door experiment with villagers and the counterpart laboratory experiment with undergraduate students are organized into three main results. The first result reveals how high cooperation rates can be observed in a

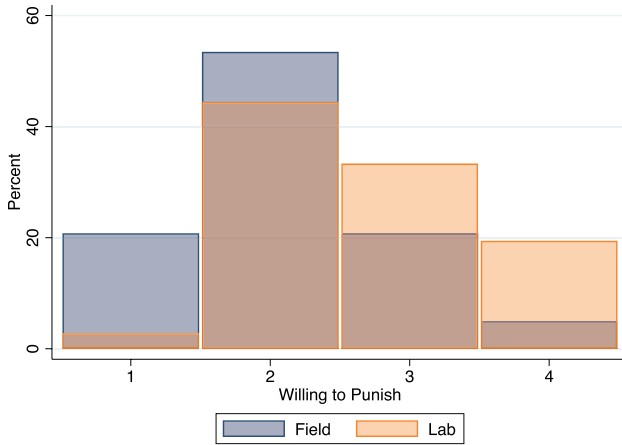


Fig. 3 Willingness to Punish

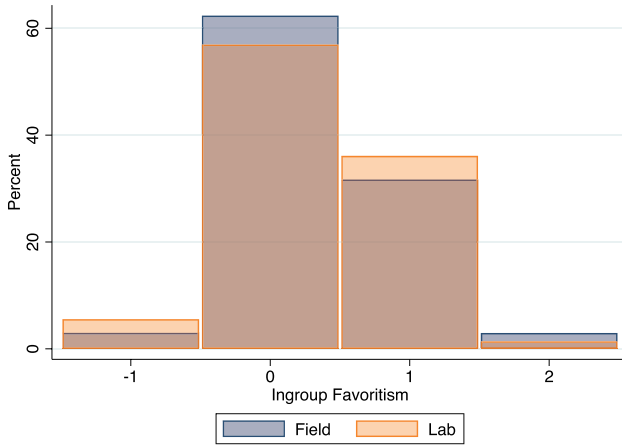


Fig. 4 In-group Favoritism

simultaneous prisoner’s dilemma experiment when villagers know the anonymous opponents in a game are their community members. The second result illustrates the impact of community membership on cooperation. The third result addresses concerns regarding the external validity of laboratory experiments with university student subjects.

Result 1 Almost full cooperation was achieved when villagers played a one-shot simultaneous prisoner’s dilemma game with their anonymous community members.

Table 3 reports the results of the door-to-door experiment with villagers. Only two of 59 villagers (3.4%) chose not to cooperate in the in-community treatment. This suggests that “Prisoner’s Dilemma” is no longer a dilemma for neighboring

Table 3 Door-to-door Field Experiment with Villagers

<i>Villagers</i>	#Send	<i>N</i>	Cooperation	S.D
Out-community	34	42	0.81	0.40
In-community	57	59	0.97	0.18
Field Total	91	101	0.90	0.30
Diff. (MW Test)	$z = -2.58, p = 0.0098$			

rural villagers in Japan. It rather seems that they easily coordinate on the efficient equilibrium in which cooperation is the established convention. At the group level, 16 out of 18 groups in the in-community treatment achieved full cooperation (i.e. the socially optimal outcome).

The cooperation rate observed in the in-community treatment is surprisingly high when compared with previous findings. Goette et al. (2006) find that 69% of Swiss Army officer trainees cooperate in a similar simultaneous prisoner's dilemma game when they interact with a member of their own platoon. However, villagers' in-community cooperation rate is more comparable to but still higher than those in a later super game of infinitely repeated prisoner's games or those in a public goods experiment with punishment opportunity (Camera et al., 2012; Dal Bó & Fréchet, 2018; Fehr & Gächter, 2000). This finding suggests that a set of beliefs and local norms of cooperation that subjects bring to the frame of an experiment can regulate behavior in the experiment.

Result 2 In the field experiment with villagers, cooperation was significantly higher in the in-community compared with the out-community treatment. The effect of community membership on cooperation was robust after controlling for the potential confounders.

Cooperation was significantly higher in the in-community than in the out-community treatment at the 1% risk level (Table 3).⁷ This treatment effect remained robust after controlling for individual characteristics as well as other potential confounders. Table 4 reports the estimation results from probit regressions, where the dependent variable equals 1 if cooperate. The coefficient of the in-community treatment is positive and statistically significant at 5% in all model specifications. Models 2 and 4 include individual characteristics (such as gender, age, and social value orientation presented in Table 2) as independent variables. The estimation results indicate that pro-social behavior is positively associated with general trust and age at the individual level.

Models 3 and 4 include potential cofounders as control variables. *Male experimenter* aims to capture the experimenter gender effect. *Cancellation* equals 1 if there

⁷ For the group-level test, seven of 12 groups in the out-community treatment achieved the socially optimal outcome. The Mann–Whitney *U* test shows a statistically significant difference at the 10% level between the in-community and out-community treatments in the number of groups that converge to full cooperation ($z = 1.906, p = 0.0567, N = 30$).

Table 4 Field data regression result (Robust standard error probit)

	Model 1	Model 2	Model 3	Model 4
	Coeff (<i>t</i> value)	Coeff (<i>t</i> value)	Coeff (<i>t</i> value)	Coeff (<i>t</i> value)
Incommunity	0.950 (2.46)**	1.151 (2.56)**	0.958 (2.47)**	1.172 (2.42)**
Male		− 0.25 (0.56)		− 0.4 (− 0.80)
Age		0.035 (2.38)**		0.041 (2.67)***
Trust		0.746 (2.34)**		0.847 (2.61)***
Reward		0.36 (0.84)		0.358 (0.87)
Punish		0.499 (1.44)		0.567 (1.46)
Gfavor		0.409 (1.33)		0.555 (1.92)*
Constant	0.876 (3.91)***	− 5.746 (− 2.90)***	0.947 (2.38)**	− 6.302 (− 2.98)***
Male experimenter			− 0.144 (− 0.43)	− 0.089 (− 0.28)
Cancellation			− 0.166 (− 0.43)	− 0.202 (− 0.45)
Village fixed effects			√	√
<i>N</i>	101	101	101	101
Pseudo <i>R</i> ²	0.11	0.23	0.11	0.25
lnLikelihood	− 29.18	− 25.08	− 28.99	− 24.62

*Denotes $p < 0.1$ **Denotes $p < 0.05$ ***Denotes $p < 0.01$

was any cancellation in the group. Since four villages were merged to form the current town in 2004, *Village fixed effects* aim to capture the potential, historical and cultural differences between the old village districts. The estimation results show that the treatment effect is robust to these concerns. Observed treatment effect of social group membership is consistent with previous findings in group identity and in-group favoritism (Bernhard et al., 2006; Chen & Li, 2009; Goette et al., 2006).⁸

⁸ Another competing explanation for the significant treatment effect might be out-group hostility (Ligon and Schechter, 2012).

Table 5 Counterpart Laboratory Experiment with University Students

<i>Students</i>	<i>#Send</i>	<i>N</i>	<i>Cooperation</i>	<i>S.D</i>
Out-classmate	9	48	0.19	0.39
In-classmate	10	24	0.42	0.50
Lab Total	19	72	0.26	0.44
Diff. (MW Test)	$z = -2.07, p = 0.0389$			

What makes this Japanese field peculiar is that a very high cooperation rate was observed even when villagers played with anonymous random out-community members. Only eight of 42 villagers (19%) chose not to cooperate in the out-community treatment. Thus, 81% of villagers still cooperated even though they were paired with any stranger in the town who were not a part of their community. This cooperation rate is fairly high given the size of the town, noting the population of about 10,000 and the area, which is ten times bigger than Manhattan. This suggests that cooperation can emerge under anonymous random matching with relatively big group size even though it is not necessary that cooperation can stabilize over time if they continue to match with strangers.

Regarding villagers' motivation for cooperation in the experiment, because we are not able to control any kind of future interactions that may or may not occur outside the frame of the experiment, I am not able to rule out a possibility of super games motivations such as punishments, rewards, and their willingness to avoid the "second-order free rider problem" (even though they can not identify their opponents in the experimental game especially in the out-community treatment). Further investigations will be required.

Result 3 Average cooperation rates were much higher among villagers than university students in all treatments. Although a big difference in the average was observed, a significant treatment effect of social group membership was preserved in both laboratory and field.

Table 5 reports the results of the counterpart laboratory experiment with university students. While 41.7% of student subjects cooperated when they were paired with anonymous classmate fellows, only 18.8% of them cooperated when they were paired with anonymous strangers. The difference is statistically significant at the 5% risk level as shown in Table 5. Both average cooperation rates and treatment effects are consistent with previous studies (Cooper et al., 1996; Goette et al., 2006).

With respect to external validity concerns, this result supports a consistent finding in the literature that university students are less pro-social compared with non-student subjects (Anderson et al., 2013; Bortolotti et al., 2015; Burks et al., 2009; Snowberg & Yariv, 2021). However, Bigoni et al. (2013) report the opposite result, and also the magnitude of the difference varies across studies depending on the type of non-student adult population. These results challenge the generalization of results from a student population regarding the magnitude of cooperation rates. Nonetheless, the good news is that a treatment effect is

Table 6 Pooled data regression result (Robust standard error probit)

	Model 5	Model 6	Model 7	Model 8
	Coeff (<i>t</i> value)	Coeff (<i>t</i> value)	Coeff (<i>t</i> value)	Coeff (<i>t</i> value)
Incommunity	0.796 (3.25)***	0.926 (3.59)***	0.677 (2.03)**	0.789 (2.24)**
Field	1.865 (7.69)***	1.887 (7.15)***	1.763 (5.75)***	1.764 (5.03)***
Incommunity × Field			0.273 (0.54)	0.306 (0.57)
Male		0.31 (1.17)		0.297 (1.11)
Trust		0.416 (2.00)**		0.416 (2.01)**
Reward		0.4 (1.56)		0.42 (1.62)
Punish		0.134 (0.83)		0.126 (0.77)
Gfavor		0.23 (1.07)		0.23 (1.08)
Constant	− 0.935 (− 4.80)***	− 4.099 (− 3.74)***	− 0.887 (− 4.23)***	− 4.084 (− 3.74)***
<i>N</i>	173	173	173	173
Pseudo <i>R</i> ²	0.39	0.44	0.39	0.44
lnLikelihood	− 68.8	− 63.57	− 68.65	− 63.41

**Denotes $p < 0.05$

***Denotes $p < 0.01$

preserved both in student subjects and non-student villagers, implying that the treatment effect of social group membership is externally valid.

The treatment effect of social group membership and a big difference in cooperation between villagers and students remained robust after controlling for individual characteristics. Table 6 reports the estimation results from pooled probit regressions, where the dependent variable equals 1 if cooperate. The coefficient of the *Incommunity* treatment is positive and statistically significant at 5% in all model specifications. This indicates that a treatment effect is preserved both in the field and laboratory. The coefficient of a *Field* dummy is positive and statistically significant at the 1% level in all model specifications. This indicates a statistically significant mean difference exists between the field and laboratory. Models 7 and 8 show that the coefficient of the interaction term between *Incommunity* and *Field* is positive but not statistically different from 0. This implies that the effect of community membership in the field is not statistically greater than the effect of classmate matching in the laboratory.

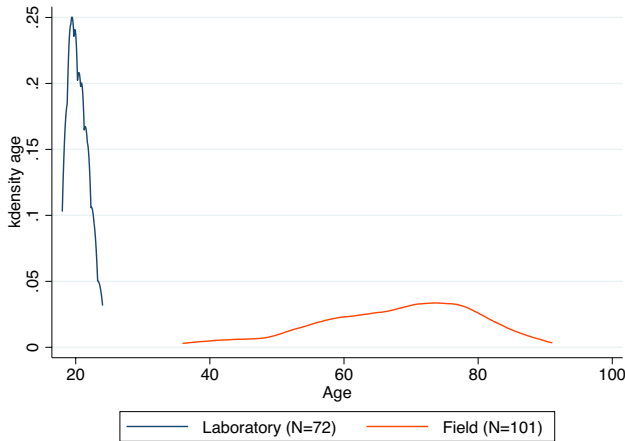


Fig. A1 Age distribution of participants

Models 6 and 8 include individual characteristics (such as gender and social value orientation presented in Table 2) as independent variables. The results indicate that pro-social behavior is positively associated with general trust at the individual level. Note that *Age* was removed from these models with pooled data because *Age* was highly correlated with whether field or laboratory. The *t*-test with unequal variances shows a statistically significant difference at the 1% level in *Age* between villagers and students ($t = -40.15$, $p < 0.001$). The youngest age in the field was 36 while the oldest age in the laboratory was 24. Figure A1 in Appendix shows the comparison of age distribution (kernel density estimates) between villagers and students. Table A1 in Appendix presents the estimation results of regressions that add *Age* to the models reported in Table 6. The coefficients of *Field* in Models 9 and 11 are still significant at the 5% level after including *Age* in the models. However, the coefficient becomes insignificant after controlling for individual social value orientation measures, shown in Models 10 and 12.

The following two observations suggest that observed full cooperation is likely to be unique to neighboring villagers in the rural area of Japan and is unlikely to be replicated with non-student samples of comparable ages in urban areas. First, Table A2 in Appendix shows cooperation rates by age groups. The result indicates that cooperation rates are stable across these age groups. The cooperation rate of the younger group is 1 or nearly 1 in the in-community treatment (1.00 for 60 years old or younger, $N = 15$; 0.97 for 70 years old or younger, $N = 31$). Second, Table A3 shows cooperation rates by age groups, which are calculated based on the percentage of choosing the cooperative strategy in prisoner's dilemma games among the representative sample of the United States population reported in Snowberg and Yariv (2021). This provides no evidence that cooperation rates among the representative sample are significantly associated with age groups. These observations suggest that age is not solely responsible for the big difference in the average cooperation rates observed between villagers and university students. Nonetheless, my data does not

Table A1 Pooled data regression result (Robust standard error probit)

	Model 9	Model 10	Model 11	Model 12
	Coeff (<i>t</i> value)	Coeff (<i>t</i> value)	Coeff (<i>t</i> value)	Coeff (<i>t</i> value)
Incommunity	0.873 (3.44)***	0.961 (3.58)***	0.781 (2.30)**	0.851 (2.34)**
Field	1.722 (2.59)***	0.816 (1.16)	1.669 (2.43)**	0.741 (1.00)
Incommunity × Field			0.204 (0.40)	0.246 (0.45)
Age	0.004 (0.29)	0.023 (1.62)	0.003 (0.25)	0.023 (1.57)
Male	0.289 (1.15)	0.272 (1.03)	0.277 (1.10)	0.259 (0.97)
Trust		0.530 (2.37)**		0.527 (2.35)**
Reward		0.414 (1.58)		0.430 (1.63)
Punish		0.191 (1.13)		0.184 (1.07)
Gfavor		0.219 (1.00)		0.220 (1.01)
Constant	− 1.250 (− 3.15)***	− 5.050 (− 4.07)***	− 1.194 (− 2.91)***	− 5.017 (− 4.03)***
<i>N</i>	173	173	173	173
Pseudo <i>R</i> ²	0.40	0.45	0.40	0.45
lnLikelihood	− 68.17	− 62.71	− 68.09	− 62.61

Denotes $p < 0.05$ *Denotes $p < 0.01$ **Table A2** Cooperation Rates by Age Groups

Age Groups	Ave Age (S.D.)	Out-Com Coop. %	<i>N</i>	In-Com Coop. %	<i>N</i>	Total Coop. %	<i>N</i>
≤ 70	58.6 (8.7)	0.81	21	0.97	31	0.90	52
> 70	77.2 (4.7)	0.81	21	0.96	28	0.90	49
≤ 60	52.8 (7.2)	0.79	14	1.00	15	0.90	29
> 60	73.6 (6.8)	0.82	28	0.95	44	0.90	72
ALL	67.6 (11.7)	0.81	42	0.97	59	0.90	101

An estimated average age for adult Japanese (Japanese over the age of 20) is 54.7, which is calculated by the author based on the Population Estimates, Statistics Bureau of Japan, www.e-stat.go.jp (last visited on June 19, 2021)

Table A3 Cooperation Rates by Age Groups among the Representative Sample in Snowberg and Yariv (2021)

Age Groups	Ave. Age (S.D.)	Cooperation Rate*	<i>N</i>
[15, 30]	23.7 (3.9)	0.56	243
[31, 50]	40.0 (5.7)	0.55	371
[51, 70]	59.8 (5.6)	0.52	346
[71, 84]	74.4 (3.5)	0.49	40
ALL	44.3 (16.0)	0.54	1000

*Cooperation rates by age groups are calculated based on the percentage of choosing the cooperative strategy in a Prisoner’s dilemma game among the Representative Sample (of the United States population) reported in Snowberg and Yariv (2021). The average δ^{SPE} (i.e. $\frac{\pi_1(D,C) - \pi_1(C,C)}{\pi_1(D,C) - \pi_1(D,D)}$) in their PD games is comparable to the one in this study. Data Source: Snowberg, Erik, and Yariv, Leeat. Data and Code for: “Testing the Waters: Behavior across Participant Pools.” Nashville, TN: American Economic Association [publisher], 2021. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2021-01-28. <https://doi.org/10.3886/E120227V1>

allow me to decompose the effects of *Field* into the effect of age and the effect of local norms shared by villagers. This is an important direction for future research.

5 Concluding remarks

A consistent body of experimental evidence of prisoner’s dilemma games has shown that experimental behaviors significantly deviate from the free-riding Nash equilibrium on average. However, so far, no experimental evidence has been documented for full cooperation. Local norms and shared beliefs in cohesive social groups regulate individual behavior in everyday economic life. Can we replicate it in a well-controlled economic experiment? In this paper, I report on a door-to-door field experiment where 120 villagers were recruited from 23 communities in a rural, mountainous Japanese village to play a simultaneous prisoner’s dilemma game. To see whether local community membership affects experimental behavior, I compared villager’s behavior under in-community and out-community random matching protocols. I also report on a counterpart laboratory experiment with 72 university student subjects to address concerns about the external validity of laboratory experiments, with the question of whether the results from a university student sample generalize to other populations.

The findings are three-fold. First, almost full cooperation was achieved when villagers played a prisoner’s dilemma game with their anonymous community members. Second, cooperation rate was significantly higher in the in-community than in the out-community treatment in the field experiment with villagers, even though a very high cooperation rate was observed even in the out-community treatment. This treatment effect of social group membership on cooperation is robust to potential confounders. Despite the lack of variation in cooperation among villagers due

to a very high cooperation rate, the pro-social behavior of villagers was positively associated with the general trust and age of participants. Third, regarding external validity concerns, a significant treatment effect of social group membership was preserved in both laboratory and field. Although, a big difference in the average cooperation rates between villagers and university students was observed. This is in line with previous studies that compared university students to representative samples (Snowberg & Yariv, 2021).

To confirm that Japanese rural villages provide substantial conditions for sustaining cooperation, it would be highly desirable to know how much of the observed difference in cooperation rates between villagers and university students is attributable to the effect of intensive social interactions among neighboring villagers. To this end, future studies should address whether and how the rate of cooperation observed among rural villagers is higher than the cooperation rate observed among non-student samples of comparable ages in urban areas. In addition, it would be important to gather further evidence from other rural areas of Japan to explore whether full cooperation is widely observed among rural villagers in Japan.

Appendix

See Fig. A1 and Tables A1, A2 and A3.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42973-021-00086-8>.

Acknowledgements This work has been funded in part by the Japan Society for the Promotion of Science (grant number 15J02390; 18H00832). I thank Kana Moriyama, Kohei Suzuki, and Nobuyuki Ito for their excellent assistance in conducting the experiments. I also thank an anonymous reviewer and participants at the 2021 International Workshop for Lab and Field Experiments for helpful comments. I greatly appreciate the hospitality of the School of Environmental Sciences at the University of East Anglia, where part of this work was done.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Alm, J., Bloomquist, K. M., & McKee, M. (2015). On the external validity of laboratory tax compliance experiments. *Economic Inquiry*, 53(2), 1170–1186.

- Anderson, J., Burks, S. V., Carpenter, J., Götte, L., Maurer, K., Nosenzo, D., Potter, R., Rocha, K., & Rustichini, A. (2013). Self-selection and variations in the laboratory measurement of other-regarding preferences across subject pools: evidence from one college student and two adult samples. *Experimental Economics*, 16(2), 170–189.
- Bernhard, H., Fehr, E., & Fischbacher, U. (2006). Group affiliation and altruistic norm enforcement. *American Economic Review*, 96, 217–221.
- Bigoni, M., Camera, G., & Casari, M. (2013). Strategies of cooperation and punishment among students and clerical workers. *Journal of Economic Behavior and Organization*, 94, 172–182.
- Bortolotti, S., Casari, M., & Pancotto, F. (2015). Norms of punishment: experiments with students and the general population. *Economic Inquiry*, 53(2), 1207–1223.
- Burks, S., Carpenter, J., & Goette, L. (2009). Performance pay and worker cooperation: evidence from an artefactual field experiment. *Journal of Economic Behavior and Organization*, 70(3), 458–469.
- Camera, G., Casari, M., & Bigoni, M. (2012). Cooperative strategies in anonymous economies: an experiment. *Games and Economic Behavior*, 75(2), 570–586.
- Cardenas, J. C., & Carpenter, J. (2008). Behavioural development economics: lessons from field labs in the developing world. *Journal of Development Studies*, 44(3), 311–338.
- Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. *American Economic Review*, 101(6), 2562–2589.
- Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1), 431–457.
- Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1996). Cooperation without reputation: experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, 218(12), 187–218.
- Dal Bó, P., & Fréchet, G. R. (2011). The evolution of cooperation in infinitely repeated games: experimental evidence. *American Economic Review*, 101(1), 411–429.
- Dal Bó, P., & Fréchet, G. R. (2018). On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature*, 56(1), 60–114.
- Duffy, J., & Ochs, J. (2009). Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior*, 66(2), 785–812.
- Efferson, C., Lalive, R., & Fehr, E. (2008). The coevolution of cultural groups and ingroup favoritism. *Science (new York, n. y.)*, 321(5897), 1844–1849.
- Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4), 980–994.
- Goette, L., Huffman, D., & Meier, S. (2006). The impact of group membership on cooperation and norm enforcement. *American Economic Review*, 96(2), 212–216.
- Goette, L., Huffman, D., & Meier, S. (2012). The impact of social ties on group interactions: evidence from minimal groups and randomly assigned real groups. *American Economic Journal: Microeconomics*, 4(1), 101–115.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, 91(2), 73–78.
- Kollock, P. (1998). Transforming social dilemmas: group identity and cooperation. *Modeling Rational and Moral Agents*, 7, 186–210.
- Ligon, E., & Schechter, L. (2012). Motives for sharing in social networks. *Journal of Development Economics*, 99(1), 13–26.
- Mitani, Y., & Shimada, H. (2021). Self-selection bias in estimating the determinants of landowners' Re-enrollment decisions in forest incentive programs. *Ecological Economics*, 188, 107109.
- Mitani, Y., Suzuki, K., Moriyama, K., & Ito, N. (2015). Describing Local Community Characteristics in Japanese Rural Villages: a community survey result and its application to explaining non-industrial private forest owners' behavior. *Natural Resource Economics Review*, 20, 85–95.
- Ruffle, B. J., & Sosis, R. (2006). Cooperation and the in-group-out-group bias: a field test on Israeli Kibbutz members and city residents. *Journal of Economic Behavior and Organization*, 60(2), 147–163.
- Snowberg, E., & Yariv, L. (2021). Testing the waters: behavior across participant pools. *American Economic Review*, 111(2), 687–719.
- Solow, J. L., & Kirkwood, N. (2002). Group identity and gender in public goods experiments. *Journal of Economic Behavior and Organization*, 48(4), 403–412.