Journal of Behavioral and Experimental Economics 100 (2022) 101915

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Economics

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

Leader and citizens participation for the environment: Experimental evidence from Eastern Europe

Tiziana Centofanti [a], Anand Murugesan [b,*,1]

[a] Department of Environmental Science and Policy, Central European University, Quellenstraße 51, 1100 Vienna, Austria
[b] Department of Public Policy, Central European University, Quellenstraße 51, 1100 Vienna, Austria

ARTICLE INFO

Keywords:
Leaders
Participation
Public goods
Environment
Lab-in-the-field experiments

ABSTRACT

We examine two aspects of decentralized governance by conducting lab-in-the-field experiments in five villages in Central and Eastern Europe. We experimentally test the effect of (1) democratically elected local leader's endorsement and (2) citizen's participation in the choice of environmental public goods (EPG) on their voluntary contributions. Our design ties the voluntary contributions in the experimental public good game to restoring a real EPG, such as a public park. We find that citizen participation in the endogenous choice of the EPG significantly increases contributions, while the leader's endorsement does not. Our results suggest that while the leader's effect is unclear, participatory decision-making can improve access to public goods and increase the beneficiaries' willingness to cost-share in rural communities.

1. Introduction

Decentralized governance and participatory decision-making have been popular political trends. Countries worldwide, including the three we study in the Central and Eastern European region, have decentralized and distributed political power to local governments (Bardhan & Mookherjee, 2006; Burki, Perry, & Dillinger, 1999; Rodríguez-Pose & Krøijer, 2009). A stated objective for the distribution of political power is improving access to public goods and services, which includes the environment of the poor in developing countries (Amel, Manning, Scott, & Koger, 2017; Ostrom, 1990; Somanathan, Prabhakar, & Mehta, 2009; Zarychta, 2020). We examine two aspects of decentralized governance by conducting lab-in-the-field experiments in five villages in Central and Eastern Europe.

We experimentally test the effect of (1) democratically elected local leader's endorsement and (2) citizen's participation in the choice of environmental public goods (EPG) on their voluntary contributions. While there is evidence suggesting a positive role of leaders in increasing citizen cooperation (Jack & Recalde, 2015; Javaid & Falk, 2015), we test the causal relationship in a natural setting with lab-in-the-field experiments by involving real elected leaders and real public goods, such as public parks and water wells.[2] We evaluate contributions to the restoration of a local environmental public good (EPG), e.g., a public park in the village, when the elected mayor endorses the restoration and when individuals have the opportunity to participate in the choice of their EPG. Our 2×2 between-subject design allows us to study the independent and interactive effects of the two treatments.

By linking individuals' contributions to the EPG in the lab to the actual field restoration of the EPG, we directly bring behavior in the lab-in-the-field experiments closer to behavior outside the lab (Barr, Packard, & Serra, 2014; Bluffstone, Dannenberg, Martinsson, Jha, & Bista, 2020). To compare subject contributions to real public goods in our study with contributions reported in other studies, we conducted the standard threshold public good game [TPGG] with a between-subject design. Any difference in contributions in the real public good game compared to giving in the standard public good game informs our understanding of individuals' choice concerning real public goods.

By engaging elected village mayors in our experiment, we add to the literature on the influence of real leaders in fostering cooperative

* Corresponding author.
E-mail addresses: centofantit@ceu.edu (T. Centofanti), murugesana@ceu.edu (A. Murugesan).

[1] We thank Maria Bigoni, Marco Casari, Geoffrey Castillo, Cristina Corduneanu-Huci, Vikram Dayal, Michael Dorsch, Subhasish Dugar, the four anonymous reviewers and the editor for comments and suggestions that significantly improved the manuscript. We thank the excellent research assistance by Iryna Koval in Ukraine, Silvia Mrázková and Lubica Stiblárová in Slovakia, and Marton Leizen and Viktor Mak in Hungary. We thank Tibor Derdak and Dr. Ambedkar School in Miskolc for tremendous institutional support against all odds in Hungary. This publication is the result of research conducted for Central European University, Private University – CEU GmbH. It was made possible by the CEU Open Access Fund and CEU ACRO research support scheme fund. Our study was approved by the Institutional Review Board at the CEU. Authors are listed in alphabetical order.

[2] See Fig. A.1 in the Appendix for an example of the Environmental Public Good (EPG), the water-spring in Porshna (Ukraine) before restoration.

https://doi.org/10.1016/j.socec.2022.101915
Received 13 August 2020; Received in revised form 4 July 2022; Accepted 5 July 2022
Available online 8 July 2022

behaviors. Kosfeld and Rustagi (2015) use real-world leaders in a public good game and found that leaders play an essential role in successfully managing collective action outcomes. While they focused on the effects of punishment by leaders and leader quality, we depart by focusing on non-coercive interventions by leaders. This departure brings other aspects of a leader's role to the fore, such as invoking social preferences, coordination, and providing information about the public good's legitimacy and value.

Others have studied the effect of engaging real-world leaders on cooperation with *artefactual* lab-in-the-field experiments (Gatiso & Vollan, 2017). Javaid and Falk (2015) engage informal leaders to observe farmers' behavior in an artefactual CPR game and Jack and Recalde (2015) use a proxy for EPGs, i.e., community members pool resources to provide environmental education material for local schools. Relatedly, del Pilar Moreno-Sánchez and Maldonado (2010) study the effect of non-coercive authorities on Colombian fishers' extraction decisions with a simulated CPR game. Their artefactual experiment used real park rangers as advisors providing information about sustainable harvesting levels. In our study, we engage the mayors similarly in a non-coercive way. However, we take another step towards external validity by combining our lab-in-the-field experiment with actual EPG restoration.

A related strand of literature on decentralized governance has examined the benefits of citizen participation in local welfare projects. People voicing their preferences through ballots in a direct democracy (Matsusaka, 2005), or active participation in the choice of projects have been documented to improve outcomes (Bardhan, 2000). Olken (2010) showed that plebiscites for deciding on local public goods increased satisfaction and knowledge about the project among Indonesian villagers. We study the effect of active democratic participation on contributions by allowing individuals to endogenously choose the EPG and compare it to contributions when the EPG is exogenously chosen (pre-identified). We implemented active participation by inviting subjects at the start of their session for a discussion about the choice of the EPG to be restored. The discussion would be held after the experimental sessions and on the day of the payments.

We conducted these experiments among rural communities in Central and Eastern Europe. Decentralized governance through the higher allocation of funds to local administrative units has been the focus of political reforms in the region over the last two decades (Péteri, 2018; Rodríguez-Pose & Krøijer, 2009). We selected five villages within a 150 km radius in Hungary, Slovakia, and Ukraine. The three countries share a long political history (Fenyvesi, 2005; Murrell, 2008), and have implemented similar decentralization policies. Our chosen villages are similar along many dimensions, including the socio-economy, demography, and cultural characteristics. Our choice of multiple sites within a geographical region was to increase the sample size, check the external validity, and generalize our findings rather than examine inter-country differences in cooperation.

We used a mechanical method to measure subject behavior in a context of low literacy and access to computers (Brañas-Garza, Cobo-Reyes, & Domínguez, 2006; Viceisza, 2016). In a choice elicitation method similar to the beans and bowls technique used by Giné, Goldberg, Silverman, and Yang (2018), we used balls and pipes to elicit choices. We considered our method suitable given the subjects' literacy level and discomfort in handwriting their choice with pencil and paper. Our method facilitated studying behavior with lab-in-the-field experiments in rural areas and marginalized communities. We adopted a finitely repeated game to allow for sufficient learning by the subjects and compare our results to related experiments (Gatiso & Vollan, 2017; Javaid & Falk, 2015).

2. Experimental design

2.1. Study locations

Fig. 1 Depicts the geographical location of the villages chosen for the study spread over Eastern Hungary, Eastern Slovakia, and Western Ukraine. The five villages are within 300 km of each other and similar in many dimensions, including the socio-economy, demography, and cultural characteristics. The economic experiments were conducted in September 2017 in Ukraine, in June 2018 in Hungary, and in September 2018 in Slovakia. We partnered with researchers at the Technical University of Košice in Slovakia and at the Uzhhorod National University in Ukraine to conduct the study. According to the criteria for preliminary screening (described in the next paragraph), our cooperators in the three regions compiled a list (containing 30 to 40 villages). We visited 6–10 villages from each region on the list to select the study villages. We conducted an unstructured interview with the mayors of each of the identified villages during the visit. We drafted a list of potential environmental public goods that could be restored from the project and then visited the potential restoration sites.

Table 1
Characteristics of villages chosen for this study.

Country-region	Village Name, Region	Population (number)	Roma People (percentage)
Ukraine West	Dovhe, Zakarpattiya	6,000	Less than 5
Ukraine NW	Porshna, Lviv	1,400	Less than 5
Hungary East	Alsózsolca, Miskolc	6,000	Around 20
Slovakia NE	Ladomirová, Presov	1,000	Around 50
Slovakia East	Zalužice, Košice	1,200	Less than 5

The criteria for selecting the final study villages were (1) a small to a medium-sized village (1000–10,000 inhabitants), (2) the presence of environmental public goods to be ecologically restored, such as a city park, schoolyard, common recreational areas, and (3) that the socio-economic and demographic characteristics of the population were not distinctive, e.g., not having a relatively high level of in or out-migration.

Table 1 presents the demographic characteristics of the chosen villages. We note the ethnic composition of the Roma minority in the study villages. The Roma population constitute about 11 million people in Europe, forming the largest minority and tracing their origins from the north of the Indian subcontinent.[3] In two of our study villages, Ladomirová (Slovakia) and Alsózsolca (Hungary), Roma people represent about 20%–50% of the village population. As ethnicity is a sensitive topic, we did not directly ask about the ethnic identity of the subject in the post-experiment survey questionnaire (Martín, Branas-Garza, Espín, Gamella, & Herrmann, 2019). Therefore we cannot estimate the exact proportion of Roma subjects in the experiments. However, based on our guess, the vast majority of the subjects in these two villages belonged to the Roma community.

2.2. Subjects

The recruitment of subjects for the experimental study was based on the village census roster. A representative random sample of households were sent the invitations to participate in the study at specified time slots during the experiments. Teams of local students distributed the invitations door-to-door in randomly selected households seeking the participation of subjects 18 years or older. As we had an inadequate turnout of randomly selected household members, further recruitment was conducted by the random street intercept method. Participation was voluntary and anonymous. Subjects were eligible if they were adults, and there were no additional requirements.[4] A total of 382 subjects participated in the lab-in-the-field experiments.

[3] Genetic studies trace their migration from Rajasthan, Haryana, and Punjab regions of modern-day India to Europe around 1500 years ago (Mendizabal et al., 2011). Roma people make about 2% and 3%–5% minority in Slovakia and Hungary, respectively.

[4] We had to decline two subjects as they were visibly not sober for the experiment.

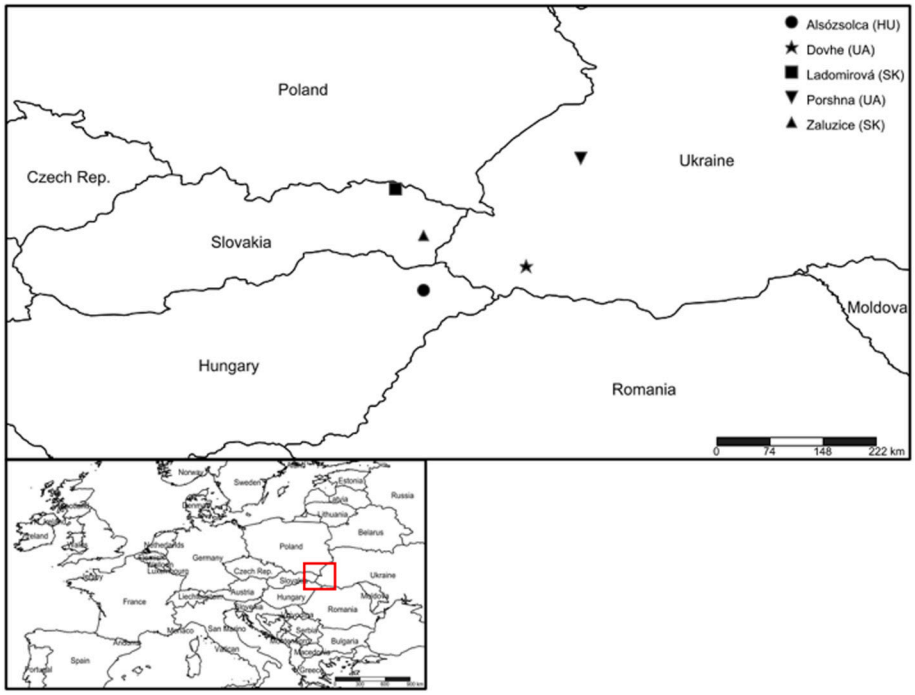

Fig. 1. Geographical location of the villages for the study.

We randomized the treatments across sessions within each village by randomly drawing the order of the sessions, with each session corresponding to a treatment. The subjects were exogenously assigned to the treatments as the treatment order was randomized. Our design addressed clustered randomization concerns that occur if individuals who know one another and share unobserved attributes are sampled.[5] Each session had between 12 to 21 subjects (multiples of 3), with the modal session containing 18 subjects. In each session, the subjects were anonymously assigned to be one subject in a group of three players. The identities of the group members were neither revealed during nor after the experiment due to the anonymity clause. As at least four groups played in the same room, the subjects could not know which of the people present were part of their group. The members of each group were randomly chosen, and the group composition remained fixed throughout the session. Table 3 presents the summary statistics, including the subject's gender, age, and educational level.

Table 2
Description and breakdown of treatments in the study.

Treatment (Description)	HUN	SLO		UKR		Sessions	Sample
	ALS	LAD	ZAL	DOV	POR		
[1] Leader endorsement [LEAD]	18	18	15	33	12	6	96
[2] Citizen participation [PART]	13	18	15	–	–	3	46
[3] Leader x Participation [LEAD_PART]	16	14	15	–	–	3	45
[4] Control [CONTROL]	18	18	18	42	36	7	132
[5] Standard PGG [TPGG]	18	18	15	–	12	4	63
Total	83	86	78	75	60	23	382

Notes: Country and village abbreviations: HUN = Hungary; SLO = Slovakia; UKR = Ukraine; ALS = Alsózsolca; LAD = Ladomirová; ZAL = Zalužice; DOV = Dovhe; POR = Porshna.

2.3. Experimental treatments

Table 2 presents the two dimensions of the treatments [1–4], i.e., the leader endorsement and citizen participation implemented with the 2 × 2 between-subject design. We report the breakdown of the sessions and the sample size in each of the five treatments. We engaged the democratically elected mayor in each village to implement the leader endorsement treatment [LEAD]. We operationalized the non-coercive endorsement of the leader by showing the subjects a short video of their mayor encouraging them to contribute to their local EPG with the following sentence:

> "*For beautifying our community, I strongly support this project and encourage you to contribute your resources to the common box. Remember that we all benefit from cooperation*".

The participation treatments were subjects contributing to (1) an endogenously chosen EPG in their village [PART] and (2) an exogenously chosen EPG to be restored [CONTROL]. We implemented the participation treatment [PART] by informing the subjects that they could participate in the choice of the EPG after the conclusion of all the sessions and when the payments have been made. Our simplified ex-post discussion in the participation treatment aimed to reduce the interactions between the subjects before the experimental sessions.

The deliberations in the community discussion would include where (type of the public good) and how to restore the EPG. In the no participation treatment [CONTROL], the EPG was chosen by the authors in consultation with the village mayor before the experimental sessions (as explained in Section 2.1). In this treatment [CONTROL], the subjects were shown a picture of the chosen EPG, for example, a local water spring as shown in Fig. A.1, to be restored in their village. Subjects could only decide how to restore that specific pre-identified EPG with the funds collected in the experiment. The subjects in the participation treatment [PART] could choose what to restore (unlike in the CONTROL) and how to restore the EPG (similar to

[5] We checked for balance of pre-treatment covariates between each of the three treatments in comparison to the control and find no significant differences (see Table A.1 in the appendix).

the CONTROL), after the experimental payments were made. In the interactive treatment [LEAD x PART], the subjects were exposed to both the leader's endorsement [LEAD] and could participate in the discussion of what to restore [PART].

To compare our results in the *real* public good games, i.e., contributions to the EPG, with those in standard public good games, we conducted an additional treatment [5] where subjects played a threshold public good game (TPGG) with the same parameters as in our "real" public good games except that they would receive their contributions back in cash (as typical in the literature on public good games Ledyard, 1995). We detail the parameters in Section 2.4 discussing the payoffs structure.

The TPGG was played in four of the five villages. The leader endorsement treatments [LEAD] were implemented in all locations. The participation treatments [PART] were implemented in Hungary and Slovakia, partly explaining the variation in sample size across treatments. We were severely restricted in recruiting additional subjects in each of these villages. Some sessions had more subjects than others; correspondingly, sample sizes in the treatments varied. However, the sample size on all the treatments were adequate to detect at least medium-sized effects with 80% power.[6]

The economic experiment consisted of 6 rounds in each session.[7] We employed the finitely repeated game to allow for learning on the part of subjects and make it comparable to closely related experiments (Gatiso & Vollan, 2017; Javaid & Falk, 2015). The repeated game allows us to measure behavior in settings where contributions are observable and temporally dynamic e.g., the behavior of conditional cooperators. For example, in many contexts, such as the maintenance of public parks, citizens can volunteer to contribute their time and efforts for the upkeep of the park, depending on the observed contributions of other members of the community (Carlsson, Johansson-Stenman, & Nam, 2014). In the context of this study, the subjects could potentially volunteer their in-kind effort in the restoration of EPGs over time. However, we do not measure the citizens' behavior after the conclusion of the experiments as it is outside the scope this study.

In each round, the subjects had to choose how much to contribute to the public fund. As standard practice, information about the average group contribution to the public good was provided to the three group members after each round.[8]

2.4. Payoff structure

The three-member public good game we implemented has the payoff structure represented in Eq. (1):

$$U = E_i - C_i + \mu \frac{\sum_{i=1}^{n} C_i}{n} \tag{1}$$

where E_i is the individual endowment and C_i is the individual's contribution.

The marginal per capita return (MPCR, represented as $\frac{\mu}{n}$) in our public good game changes by a step at a particular threshold, i.e., when the average group contribution equals to or greater than 75% of the endowment. The design makes our game similar to a threshold public good game, where the MPCR changes at some threshold. Specifically in our design, the MPCR is 0.66 (or $\mu = 2$) if the group's contribution to the public account is less than 75% of their total endowments and; the MPCR is 1 ($\mu = 3$), if the group members' average contribution is equal to or more than 75% of their endowment to the public account.

When subjects do not expect the threshold of 75% to be reached, the Nash equilibrium in our game is the same as in the linear public good game: each individual has an incentive to free-ride, and the dominant strategy equilibrium is to contribute zero to the public good. In a one-shot public goods game where the MPCR changes to 1 at the threshold, there is a set of efficient Nash equilibria depending on beliefs about other members' contributions (Schram, Offerman, & Sonnemans, 2008). Our design with repeated contributions allowed subjects to update their beliefs. In the standard threshold public good games, the public good is not provided if the threshold is not reached (Croson & Marks, 2000). In our threshold public goods treatment [TPGG], the public good is provided whether the threshold is reached or not; but the MPCR changes at the threshold.

$$U = E_i - C_i + R(\mu \frac{\sum_{i=1}^{n} C_i}{n}) \tag{2}$$

As the primary focus of our study was examining behavior in the real public good games, we retained the same payoff structure as in the standard public good game described above for experimental control. However, the contributions in the real public good games would be used for developing and restoring the local EPG and not returned as cash to the subjects as in the standard public good game. Therefore the utility depends on the individual's valuation of the amenity (EPG), represented by R in Eq. (2). The subjects were informed they would be paid in cash, the equivalent of the balls they kept for themselves after the sessions and that the multiplier (μ) for the public good would be based on a randomly selected group's average contribution to the EPG at the end of the sessions in each village.[9]

The design with two multipliers rewarded high contributions to restore the EPG. Note that the two multipliers do not necessarily change the incentives to free-ride in our experiments. We highlight that in our setting, the contributions to the public goods in the EPG treatments were not returned to the experimental subjects as cash but directed towards the restoration of the EPG. Therefore we operationalized the design with a high multiplier and a relatively high endowment (Isaac & Walker, 1988). Our high multiplier increased the public benefits while the high relative value of the endowment was implemented to make the trade-off between public benefits and private costs comparable with similar studies (Jack & Recalde, 2015; Kosfeld & Rustagi, 2015).

2.5. Hypothesis

Although not conclusive, the experimental literature broadly indicates a positive effect of leaders on welfare outcomes e.g., by targeting subsidies or coordinating on maintenance of public goods (Basurto, Dupas, & Robinson, 2019; Jack & Recalde, 2015; Kosfeld & Rustagi, 2015). Other experimental studies strongly suggest that democratic participation increases cooperation (Dal Bó, Foster, & Putterman, 2010; Gatiso & Vollan, 2017) and improves welfare (Beath, Christia, & Enikolopov, 2017; Olken, 2010). While we were agnostic about the direction of the effect of leader endorsement [LEAD] due to our priors about citizens' distrust in elected leaders in Central and Eastern Europe (Kostadinova, 2009), we expected a positive effect of allowing citizens to participate [PART] in the choice of the public good. However, we were unable to be specific about the size of the participation effect.

[6] We use G*Power to estimate the minimal detected effect sizes (d) with 80% power and where $\alpha = 0.1$ is the probability of rejecting the null hypothesis when it is true (type I error). For a t-test of difference in means and assuming samples are equally divided between the treatment and control groups, our smallest sample on treatment (45) would allow us to detect an effect size of 0.45. With $\alpha = 0.05$, our minimal detected effect size is 0.5, which is the convention for medium effect size (Cohen 1988). The estimated effect size on our participation treatment ranges from 0.5 to 0.6 depending on the model.

[7] Subjects made decisions over six rounds in Slovakia and Hungary and, nine rounds in Ukraine.

[8] The group size in public good games varies (typically 2 to 4, but even up to 1000 Pereda et al., 2019). Its choice depends on the relevance to the setting. We chose to have three-member groups, rather than four or more, to increase the group-level variation in average contributions for analytical purposes.

[9] The random group was picked from a bowl during the participatory session after the conclusion of all the experimental sessions usually by the local senior citizen present in the room.

Fig. 2. Photos of the portable experimental lab.

We examine the combined effect of leader endorsement of the EPG and active citizen participation in the choice of the EPG. As we could not *a priori* sign the effect on the leader endorsement treatment, we were agnostic about the direction and sign of the effect in our combined treatment [LEAD x PART], where we interact both the leader endorsement and participation treatments.

Finally, we expected higher contributions in the real public good games, where contributions are used to restore the EPG, compared to the standard public good games [TPGG], where cash from the public pot is returned as per the MPCR. Our priors were based on a pilot study and due to the "realness" of the EPG compared to returns in the standard TPGG.[10]

2.6. Experimental procedure

The study was conducted in rural communities with limited or no access to computer labs and low facility in using papers and pencils.[11] Therefore we developed a mechanical method to elicit contributions in the experiments, unlike the standard procedure of using computers or pencil and paper.

A portable mobile lab (Fig. 2) was built with opaque cardboard and pipes, and we used styrofoam balls to measure subjects' choices in the public good games. Our mechanical method retains the fundamental economic incentives in experiments as the styrofoam balls represent cash while allowing subjects in low literacy contexts to intuitively understand the experimental game and to discreetly drop their contributions into the pipes. Viceisza (2016) highlighted the challenges of conducting lab-in-the-field experiments in low literacy contexts. Conducting experiments in rural areas of developing countries often necessitates that subjects be paired with assistants if using the pencil and paper method or computers to elicit choices, limiting the privacy and confidentiality of subject choices. We were able to mitigate such privacy concerns with our intuitive mechanical method.

Our use of balls, with pipes as contribution outlets is similar to (Giné et al., 2018) where subjects allocated beans, representing cash, into two different bowls representing immediate and future consumption. When compared to our method, the pencil and paper method would have been relatively more intrusive and costly as the average subject was expected to require assistance noting their responses on paper; the issue would have further amplified since we measured multiple rounds of contributions. Besides (Giné et al., 2018), other studies have used physical objects for eliciting choices. For example, Apicella, Marlowe, Fowler, and Christakis (2012) used honey sticks in public good games conducted with hunter–gatherer subjects in Tanzania, and Tognetti, Berticat, Raymond, and Faurie (2012) used rice as a continuous measure of giving in public good games conducted in rural Senegal.

During the experimental sessions, subjects were seated in the same room, each at a table separated from their neighbors' seats by a cardboard frame about one meter on either side forming a cubicle similar to an election booth (see the right panel in Fig. 2). An opaque PVC pipe (dark gray-colored, 6 cm diameter) was firmly fixed to each table, and the subjects could easily access one end of the pipe at arm's length. The shields in the booths and pipe locations were designed to protect subjects' choices from being observed by others. Each subject was given ten plastic bags: one bag for the trial round followed by nine bags for the experiment. Each bag contained ten styrofoam balls (2 cm in diameter) and marble ball. The movement of the styrofoam ball across the pipes is perceptibly noiseless, while the heavier marble ball moving along the pipe will be audible in a quiet room.

Each styrofoam ball was equal to 10 (euro) cents (or its equivalent in Ukraine and Hungary); hence each bag containing ten balls was equal to €1. In total, the nine bags for each subject corresponds to €9. In addition to this amount, each subject received €1 as a fixed participation fee. To decide, the subject had to insert the number of styrofoam balls (representing a money equivalent) that they chose to contribute into the pipe. To indicate that their decision was final, they

[10] In pilot experiments with students at the Central European University Budapest we compared contributions in a "real" public goods game to contributions in a standard public goods game. In the real public goods game, the contributions went towards the provisioning of coffee machines, toasters, etc., for the student common room. In contrast, the contributions in the classic public good game would be returned as cash. The contributions to the in-kind ("real") public good games were slightly higher than the contributions in the standard public good game in the pilot study.

[11] Even though the average self-reported education level was around middle school, some subjects were not comfortable with filling in the surveys.

inserted the marble ball into the pipe after the styrofoam balls in each round. The marble ball also served the purpose of pushing all the styrofoam balls through the pipe.

At the end of each round, the styrofoam balls from each pipe were collected in numbered plastic bags. Once all subjects had inserted the marble balls, the plastic bags containing the balls were collected, manually counted, and recorded. Subjects received information about the contribution of their respective groups via handwritten and folded paper slips after each round. The empty plastic bags were then placed at the end of each pipe, and the subjects were asked to play the next round. The process of collecting and counting the balls after each round was between 3 to 5 min, and the total length of each session was between 60 to 70 min. The contribution rounds were followed by a pencil and paper survey.

At the beginning of each session, the local researcher instructed the subjects about the rules of the game and guided them about decision-making in the experimental lab in their local language (see experimental protocols in the Appendix). The subjects then received a comprehension test related to contributions, e.g., about the marginal and collective returns from the public good game to ascertain their understanding. On the survey, a majority (80%) self-reported understanding of the instructions in the experiment (see Table 3 for descriptive statistics), which was similar to the performance in the comprehension test at the start of the experiment. We control for subjects' understanding in our main estimations.

We held a trial run to familiarize the subjects make their contributions. The trial run where the subjects familiarized themselves with the public goods game by dropping balls into the pipes was unanchored, i.e., we did not specify how much they should contribute. After the trial run (round zero), the subjects were given another opportunity to ask clarificatory questions, followed by the experiment. No communication between subjects was allowed during the experiment. The subjects were informed that they would be paid in cash after the conclusion of all the sessions and according to their choices, i.e., the balls they keep for themselves (and not dropped in the public pot). The expected game payoffs correspond to the average daily income of the subjects in Hungary and Slovakia but are higher than the daily wage in Ukraine (Kocher, Martinsson, & Visser, 2008). The significant monetary stakes make the trade-offs salient between privately keeping the money or contributing it for public good restoration and could bring the experiment closer to behavior in real-life situations.

At the end of the experiment, we collected responses of the subjects in an unincentivized survey where we measured their beliefs, general level of trust in institutions, and opinions towards environmental public goods in addition to attributes such as age, gender, education level, and self-reported level of comprehension of the experimental instructions. The surveys were not completely filled by all the subjects due to time constraints and other issues (such as limited attention or inadequacy in responding to pencil and paper-based survey questions). The surveys and experimental instructions were conducted in the local language and back-translated (to English) to verify clarity and consistency in all the study locations.

To conclude, we add a note on the restoration of the EPG in the study villages after the completion of the experiments and subject payments in the study. The contribution in the lab-in-the-field was used to restore the chosen EPG, thus realizing the environmental benefits of their voluntary contributions. After the participatory community consultation, all the villages have implemented and completed the restoration project, voluntarily contributing materials and labor. In some cases the village amplified the collective contributions in the public pot for restoration by their own efforts but documenting this was outside the scope of this study.

Table 3
Summary statistics and subject characteristics.

Variable	Obs	Mean	Std. Dev.	Min	Max
Average contribution (percentage)[a]	382	59.32	32.48	0	100
Group contribution over 75% (round, dummy)	382	0.30	0.46	0	1
Group contribution over 75% (session, dummy)	382	0.20	0.40	0	1
Understand instructions[b] (dummy)	332	0.80	0.40	0	1
Female (dummy)[b]	363	0.62	0.48	0	1
Age (in years)[b]	290	41.40	14.86	18	76
Education level[b]	373	3.56	1.33	1	6
Average contribution to EPG (percentage)	319	60.23	32.97	0	100
Average contribution in the TPGG (percentage)	63	54.75	29.54	0	100

[a]Percentage of the balls contributed (out of ten in the endowment) per round.
[b]Gender, age and education level (1 = no education, 2 = elementary school, 3 = secondary school, 4 = college, 5 = university, 6 = post-graduate and higher), understand instructions have missing observations as some survey's were incomplete.

3. Results

3.1. Descriptive statistics

Table 3 presents descriptive statistics of the subjects and average contributions in all the rounds. A total of 382 subjects participated in the lab-in-the-field experiments in the five villages.

The experiments had a higher proportion of female subjects (62%), and the average age in the sessions was 41 years. The average self-reported education level (where 1 = no schooling and 6 = post-graduate study and above) was above secondary school. In two of the five villages, Alsózsolca (Hungary) and Ladomirová (Slovakia), Roma members represent approximately 20% and 50% of the village population, respectively (as reported in Table 1). As noted before, based on our guess, the vast majority of the subjects in these two villages belonged to the Roma community. 80% of the subjects self-reported that they understood the instructions. 44% reported fully understanding, and 36% reported understanding most of the instructions. The treatments were randomized at the session level, and there was almost no attrition (except one) in the experiments.[12] The average contribution was about 59 percentage, i.e., 5.9 balls were contributed on average from the endowment of ten balls across rounds. 30% of the groups contributed over 75% (the threshold for multiplier change) of their endowment to the public good on average in the rounds. 20% of the groups contributed over 75% of their endowment to the public good at the session-level. The average contribution to the EPG was 60.23%. In the threshold public good game [TPGG] the average contribution was almost 55%, which is similar to the results with a comparable 75% threshold treatment in Cadsby and Maynes (1999) and within the range of related threshold public good games with continuous contributions (Croson & Marks, 2000; Kaczan, Pfaff, Rodriguez, & Shapiro-Garza, 2017).

We present our results with descriptive graphs followed by the estimates of the treatment effects in the regression analysis.

3.2. Descriptive results

Fig. 3 graphically presents the average contributions across treatments. On the left panel, we compare the simple average of contributions in the four treatments of the real public good games with the contributions in the standard threshold public good game. The bars on the four treatments [LEAD, PART, LEAD x PART, and CONTROL]

[12] As our subjects were randomized at the session level and we do not collect pre-treatment baseline data, we do not present means of the covariates (Mutz, Pemantle, & Pham, 2019). We find no statistical differences in subjects age, gender and education level across treatments (see Table A.1). We control for the covariates measured in the post-experiment survey to improve the precision of our main results in robustness checks (see Table A.2).

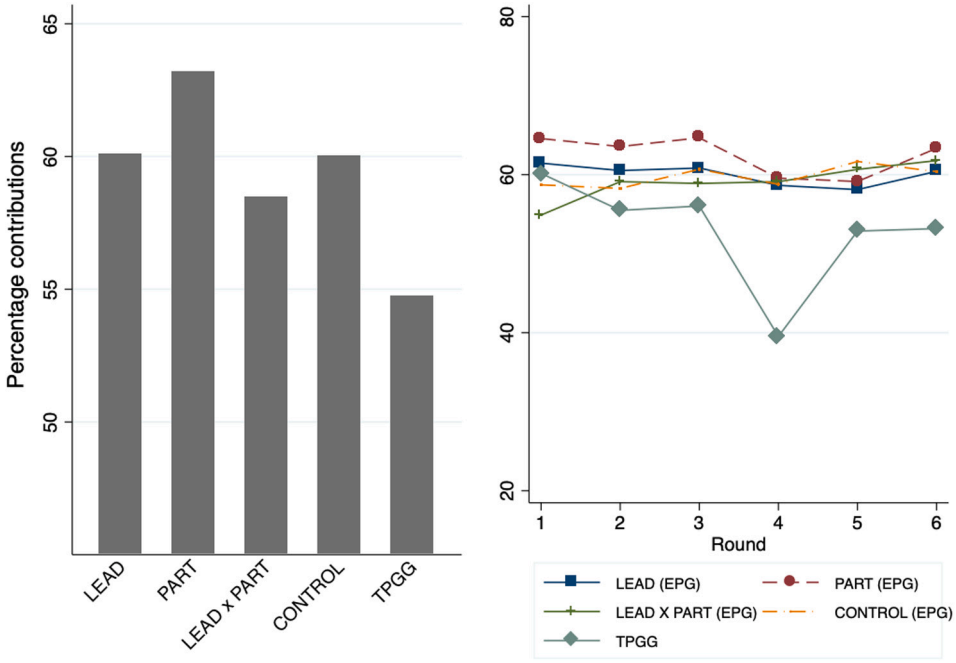

Fig. 3. Average contribution across treatments (left panel) and over the rounds (right panel).

are higher than the TPGG treatment. On average, the contributions in the real public good games (EPG treatments) were 60% of the endowment compared to about 55% in the TPGG treatment. Within the four EPG treatments, the participation treatment had the highest contributions. We present the contributions across the rounds on the right panel. We note that Fig. 3 descriptively presents unconditional means of the contributions, while the regression results presented in the next section provide estimates of the conditional means of contributions across treatments.

3.3. Regression results: Leader and participation effects

Table 4 shows the results of the contribution to the EPG in the treatments compared to the baseline treatment [CONTROL]. The dependent variable was the subjects' contributions to the public good game (balls per round, one ball = 10 euro cents). Model 1 (col. 1 in Table 4) reports the basic specification with village dummies (the baseline village is Alsózsolca, Hungary). The dummy variables at the village level control for time-invariant regional fixed effects. As we use contributions in each round as the dependent variable for our main estimates (Table 4), we include round dummies in all the models. The standard errors in all the reported estimations are conservatively clustered at the session level, which is the level of the randomized treatment (Abadie, Athey, Imbens, & Wooldridge, 2017).

We find a significant difference between the baseline [CONTROL] and the two treatments that included participation, i.e., only participation [PART] and leader and participation [LEAD_PART], in all the models. In model 1, the leader endorsement treatment [LEAD] is not significantly different from the baseline treatment. In contrast, the participation in the choice of the EPG treatment [PART] increases the contribution by about 30.7 percent (at a 5% level of statistical significance).

The interaction treatment [LEAD_PART], i.e., the mayor's endorsement of the local EPG combined with citizen participation in the choice of the EPG, increases the subjects' contribution by over 25 percent (at 1% statistical significance). We do not find a statistically significant difference between the coefficients on the citizen participation treatment and the interaction treatment where leader endorsement is combined with participation. However, we find a statistical difference between the interaction treatment and the leader treatment. We interpret the interaction treatment effect as that arising from participation rather than the leader endorsement.

In model 2 (col. 2 in Table 4), we control for higher than 75% contributions in the previous rounds at the group level by including a dummy indicator variable. The coefficient on this indicator variable is positive

Table 4
Regression results of contributions to EPG across treatments.

	Individual contribution per round		
	(1)	(2)	(3)
Leader endorsement LEAD	0.447 (0.384)	0.446 (0.376)	0.475 (0.368)
Participation PART	1.847** (0.837)	1.835** (0.829)	1.850** (0.832)
Leader x Participation LEAD_PART	1.507*** (0.387)	1.493*** (0.387)	1.523*** (0.387)
Ladomirová (SK) (Village dummy)	0.569 (0.651)	0.559 (0.647)	0.575 (0.639)
Zalužice (SK) (Village dummy)	5.569*** (0.549)	5.376*** (0.551)	5.304*** (0.533)
Dovhe (UA) (Village dummy)	4.378*** (0.604)	4.314*** (0.600)	4.313*** (0.585)
Porshna (UA) (Village dummy)	5.093*** (0.571)	4.958*** (0.578)	5.024*** (0.579)
Previous round over 75 percent		0.225* (0.131)	0.221* (0.132)
Understand instructions			0.386* (0.209)
Round dummies	Yes	Yes	Yes
Constant	2.644*** (0.559)	2.350*** (0.521)	2.168*** (0.541)
Observations × rounds	2397	2397	2397
Subjects	318	318	318

OLS regression with the dependent variable: Individual contributions per round [0,10]. Independent variables on the treatments are all binary variables. Model 1 shows 30.7% increase with PART and 25% increase with LEAD_PART; Models 2 and 3 show similar order of treatment effects; Standard errors (in parentheses) clustered at session level; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

as expected and significant at 10% level. The size and significance of the coefficients on our treatment variables remain similar to model 1. The coefficient on the dummy indicating that the average contributions of the group in the previous round was at least or exceed 75% of the endowment is positive and significant effect, as expected. In Model 3 (col.3 in Table 4), we include a dummy indicating a self-reported understanding of instructions in the experimental sessions. The coefficient on the dummy is positive and significant at 10% level, suggesting that people self-reporting understanding the instructions contributed about 5% higher on average across treatments. There were no statistical differences in the self-reported understanding of instructions across treatments.[13] The size of the coefficients on our treatment variables increases marginally, but the significance remains unchanged.

Alsózsolca (Hungary), the baseline village, and Ladomirová (Slovakia) had predominantly Roma subjects in the experiment. This may explain the similar contributions in Ladomirová and the baseline village, i.e., no statistical difference.

3.3.1. Other controls and robustness checks

We collected data on subject characteristics and reported trust levels in an unincentivized survey at the end of the experiment. We controlled for these survey measures, i.e., age, gender, and education level, and reported levels of trust of the subjects to improve the precision of estimates (see the Appendix, col. 1 in Table A.2). The size and statistical significance of the participation treatment in our preferred specification (col. 3 of Table 4) increase (from 30.7% to 38%) with the inclusion of these control variables. Age and gender are not significantly associated with giving in the experiments, while the education level and trust in other villagers is positively associated with voluntary contributions to local public goods. The sample size decreases with each additional control due to missing survey data, but the results are consistent for the treatments. We check the robustness of the estimates on the sub-sample of subjects who self-reported fully understanding the instructions (col. 2 in Table A.2).

As another robustness check of the results, we estimated our preferred specification (col. 3 in Table 4) with contributions in only the first round as the dependent variable (col. 3 in Table A.2). The contribution in the first round is devoid of response to other members' contributions. The results are similar to the main estimates except that the participation treatment [PART] effect is marginally higher (33%) and the interaction effect (19%) is lower than in our preferred specification.

4. Discussion

As countries worldwide decentralize governance by empowering local elected leaders with resources and decision-making powers, it is necessary to examine its effectiveness causally.

Our lab-in-the-field study was designed to examine two facets of decentralization: (1) citizen participation in resource allocation i.e., choice of the public good to be restored and (2) non-coercive influence of democratically elected local leaders. We find that voluntary contribution to public goods is significantly higher (about 30%) when subjects have the opportunity to choose it. The combined influence of the local mayor's endorsement and endogenous choice also results in a significantly higher contribution to the public good. The leader's endorsement of the public good without participation does not significantly increase contribution.

The results of the participation treatment are in line with previous studies finding a positive effect of direct democracy in development projects by matching resource allocation closely with citizen preferences (Beath et al., 2017; Gatiso & Vollan, 2017; Olken, 2010). Our results enrich the findings of previous studies showing a positive effect of participation on contributions to environmental public goods in artefactual lab-in-the-field experiments (Gatiso, Vollan, & Nuppenau, 2015) or by proxy for environmental goods (Jack & Recalde, 2015). We contribute to the literature by tying the voluntary contributions in the public good game with the restoration of environmental public goods under participatory systems. (Rodríguez-Pose & Krøijer, 2009) find that sub-national governments with their own revenue source through local taxation are better able to respond to local demands and promote greater efficiency. Our finding of higher contributions when citizens in rural communities can participate in the choice of the public goods suggests a possible mechanism underlying the efficiency of local taxation in rural contexts.

We add to the literature studying the role of leaders by operationalizing an endorsement from a locally elected leader rather than an experimentally assigned leader (d'Adda, 2017; Gächter, Nosenzo, Renner, & Sefton, 2012) or non-democratically elected leaders (del Pilar Moreno-Sánchez & Maldonado, 2010; Javaid & Falk, 2015). While this design brings realism into the experiment, more so as the contributions go to restoring a real public good resembling a natural decision-making setting, it may be argued that it suffers from endogeneity issues such as sample selection bias. For example, cooperative mayors are more likely to allow such experimental studies to be conducted in their villages (Corduneanu-Huci, Dorsch, & Maarek, 2021; Das, 2020). However, our qualitative evaluation of mayors in the studied villages records that both engaged and disengaged leaders hosted the study.

In contrast to several other studies (Gatiso & Vollan, 2017; Jack & Recalde, 2015; Kosfeld & Rustagi, 2015), we do not find a statistically significant effect of leaders' influence on cooperation. This result could be due to the variation in our leadership treatment from these studies, i.e., our leaders were democratically elected, and our treatment involved a non-coercive endorsement. Our result of an insignificant leader effect is similar to Javaid and Falk (2015). However, their study engaged traditional leaders in Pakistan, while we engaged democratically elected leaders. In our study, the level of subjects' trust in their mayor may partly explain the insignificant effect of the leader's endorsement on contributions to their local EPG. In the two villages where the Roma population was substantial but the mayor was non-Roma, the contributions were significantly lower than in other villages. Studies on leader endorsements indicate that endorsements from ethnic in-group leaders should increase support for policies and that endorsements from ethnic out-group leaders could even decrease support for policies (Haas & Khadka, 2020).

Our study incidentally contributes to the nascent literature on the Roma minority in Europe (Martín et al., 2019). Many factors may explain the significantly lower contribution of subjects in the Roma villages in our study. Their preference for environmental public goods may be secondary to other impending needs, e.g., nutrition, health care, education, as they are severely socially and economically disadvantaged (Harper, Steger, & Filčák, 2009). Our results contrast (Martinsson, Villegas-Palacio, & Wollbrant, 2015) findings on the role of social status on cooperative behavior but are similar to d'Adda (2017), where they also find that lower social status citizens give less.

The high contributions to the actual public goods in our study, approximately 60% on average to EPG, raises the question about contributions to real public goods compared to standard public good games. This higher contribution to EPG is even more surprising given the uncertainty of returns since several steps are involved before the public good is realized (or restored) relative to returns in the case of classic public good games (Levati & Morone, 2013). Furthermore, contributions to EPG would be affected by preferences for the environment, which may differ across individuals.

The high contributions in the participation treatment could have been influenced by the prospect of face-to-face interactions with other

[13] We report the regressions on the sub-sample of those self-reporting understanding the instructions in the Appendix. The results of the treatments are qualitatively similar, but the size of the treatment effect increases.

Table A.1
Balance across treatments.

	(1) LEAD	(2) PART	(3) LEAD x PART
Age	−0.002	0.003	0.003
	(0.004)	(0.003)	(0.005)
Gender	0.0915	−0.101	0.0713
	(0.133)	(0.116)	(0.0891)
Education	−0.094	−0.021	−0.034
	(0.067)	(0.076)	(0.068)
Constant	0.778*	0.247	0.247
	(0.392)	(0.395)	(0.445)
Observations	1958	1518	1551

Regression results with Treatment dummy as dependent variable (compared to control). Standard errors (in parentheses) clustered at session level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

citizens during the ex-post discussion on the choice of public goods. However, the voluntary participatory discussions by design were held after the payments were made. Despite the high contributions, less than 15 people per village participated in these discussions. This suggests that the participation effect may be similar to the democratic dividend found in the literature: that democratic institutions may affect cooperative behavior directly in addition to instrumental effects due to incentives (Dal Bó et al., 2010; Markussen, Putterman, & Tyran, 2014). Our results suggest that further research is needed to identify differences in giving between standard public good games and real public good games (such as the EPG in this study) as the literature relies on such games to infer behavior in real situations (Camerer, 2015; Reindl, Hoffmann, & Kittel, 2019). Standard public good games could capture the average giving in many contexts but may incorrectly estimate giving in settings where intrinsic motives are triggered, such as democratic participation and environmental issues.

5. Conclusions

This study contributes to the literature on causally measuring the role of decentralized governance in improving access and the quality of public goods. We analyzed two aspects of decentralization in public good creation: the non-coercive role of local elected leaders in influencing choices and, the effect of democratic participation in the choice of the public good. To the best of our knowledge, our study is the first to combine a lab-in-the-field experiment with the participatory restoration of real public goods *in situ*. Of note is that the high contributions in the EPG spurred cooperation among the residents that extended beyond the research purpose of this study. The fund obtained to restore the public good was used as a seed-money in three villages to attract and obtain further resources.

Despite the limitations, our study design takes steps to bridge the gap between behavior in the lab and in the field. We conducted the lab-in-the-field experiments with non-standard subjects in multiple locations, engaged real elected leaders as part of our experimental treatments, and linked voluntary contributions in the lab public good game to the actual field restoration of local environmental public goods. The portable lab design accommodated subjects with low literacy levels and facility in using computers, thus allowing representative local populations to participate in the experiment. Such mechanical methods can be developed further to conduct studies in areas that are less accessible.

Appendix

See Fig. A.1, Tables A.1 and A.2.

Table A.2
Robustness checks of main regression estimates.

	Voluntary contribution		
	All rounds		Round 1
	(1)	(2)	(3)
Leader endorsement	0.316	0.591	0.672
	(0.353)	(0.378)	(0.408)
Participation	2.298**	2.381**	2.082**
	(0.981)	(0.967)	(0.760)
Leader x Participation	1.411***	1.490***	1.145***
	(0.434)	(0.411)	(0.393)
Ladomirová (SK)	−0.368	−0.403	0.775
	(0.490)	(0.447)	(0.473)
Zalužice (SK)	4.337***	4.290***	5.442***
	(0.633)	(0.590)	(0.471)
Dovhe (UA)	3.322***	3.451***	4.536***
	(0.545)	(0.452)	(0.598)
Porshna (UA)	4.105***	5.164***	4.731***
	(0.483)	(0.400)	(0.487)
Previous round over 75 percent	0.256	0.193	n.a.
	(0.158)	(0.123)	
Trust in other villagers	0.537***	0.445**	–
	(0.127)	(0.184)	
Understand instructions	0.212	n.a.	0.223
	(0.254)		(0.308)
Other controls	Yes	Yes	–
Round dummies	Yes	Yes	n.a.
Constant	0.799	0.875	2.190***
	(0.777)	(0.760)	(0.478)
Observations	1781	1483	319

Columns 1 and 2 include other controls: age, gender, education and, self-reported trust in other villagers. Column 2 is on the sub-sample who reported fully understanding instructions. Column 3 with contributions in the first round only. Standard errors clustered at session level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. n.a. signifies not applicable.

Instructions of the experiment

Protocol treatment 1 (Leader/Identified public good)

Introduction

We would like to thank you for accepting this invitation. In the following exercise we kindly ask you to participate in a project about individual decisions and public goods in your village (environmental restoration (beautification of the area, such as addition of benches, trees, bushes, etc. as in this picture) [*show picture of the area*].

You will spend about 1.5 h here: making choices that will determine your payments and answering a short survey at the end. The funds for this project have been provided by Central European University Budapest.

We will use the registration number you obtained to match your choices with the payments to maintain your participation anonymous.

You will receive vouchers corresponding to 1 euro as a participation fee and anything else you earn today will be added to this amount and paid to you in vouchers.

Besides this exercise you will be asked to voluntarily participate in a consultation to decide the type of improvements in the public area to be restored [*Show picture of the area*]. You will decide the type of improvements that can consist of planting trees and bushes, installing benches, and creating paved paths, etc. The consultations and payments will be held // (TBA) // in this location. Participation is voluntary.

[show the slide of the mayor's picture with statement].

[Mayor's statement: "**for beautifying our community I strongly support this project. Remember that we will all benefit from cooperation**"]

Fig. A.1. Water-spring identified to be restored (Porshna, Ukraine).

Your contributions to fund the restoration project will be managed by the mayor.

Methodology

During this exercise, we will use styrofoam balls to calculate your payments. [*Show the ball and tell them that the balls are inside the bags on the desk*]

10 cents (= 1 styrofoam **ball)**

[*Show them a styrofoam ball and tell them that it is equal to 10 cents per ball*]

Every player has several bags on the desk and each bag contains 10 styrofoam balls.

You will play multiple rounds. (Number of rounds: 1 practice round, 9 rounds total for payments of which 6 rounds for the treatments with contribution to the identified or unidentified public good in the village.†

Each round you have to make a decision about how much you want to contribute to the common box. When you are asked to make the decision, place the number of styrofoam balls you want to contribute to the common box in the pipe. When you are done with inserting the styrofoam balls in the pipe, insert the marble [*show the marble ball*] to indicate that you have ended the round. You will be playing in groups of three, but you will not know who are the members of your group.

Your contributions into the common box will be multiplied by 2 as a matching contribution by the project funds. If your contribution to the common box is equal or above 75% of total, this amount will be multiplied by 3 (matched by the project funds).

[*Point to the slide about the multiplier effect*]

Every round the styrofoam balls that you have kept for yourselves should be inserted in the paper bag placed on the table. At each new round use a new bag of 10 styrofoam balls. All the members will receive exactly the amount they kept for themselves (all balls in the paper bags) whereas the contribution to the common box (with corresponding matching) will be used for the restoration of the common area.

Comprehension test

Finally, to get ready for this exercise, please let us know if you have difficulties reading or writing numbers. Also keep in mind that from now on you should make your decisions individually and should not talk or discuss with others until the end of this exercise.

Please answer the question in the comprehension test on your desk. (show them a copy of the comprehension test)

We will first have one round of practice that will NOT count for the real payments (just for practicing).

Practice round with balls.

† We discuss the results from the first six rounds for the local environmental public good. We implemented a distant environmental public good check in the last three rounds in Hungary and Slovakia, but do not discuss it in this paper.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). *When should you adjust standard errors for clustering?*: *Technical report*, National Bureau of Economic Research.

Amel, E., Manning, C., Scott, B., & Koger, S. (2017). Beyond the roots of human inaction: Fostering collective effort toward ecosystem conservation. *Science*, *356*(6335), 275–279.

Apicella, C. L., Marlowe, F. W., Fowler, J. H., & Christakis, N. A. (2012). Social networks and cooperation in hunter-gatherers. *Nature*, *481*(7382), 497–501.

Bardhan, P. (2000). Irrigation and cooperation: An empirical analysis of 48 irrigation communities in south India. *Economic Development and Cultural Change*, *48*(4), 847–865.

Bardhan, P., & Mookherjee, D. (2006). *Decentralization and local governance in developing countries: a comparative perspective*. Mit Press.

Barr, A., Packard, T., & Serra, D. (2014). Participatory accountability and collective action: Experimental evidence from albania. *European Economic Review*, *68*, 250–269.

Basurto, M. P., Dupas, P., & Robinson, J. (2019). Decentralization and efficiency of subsidy targeting: Evidence from chiefs in rural Malawi. *Journal of Public Economics*, 4047.

Beath, A., Christia, F., & Enikolopov, R. (2017). Direct democracy and resource allocation: Experimental evidence from afghanistan. *Journal of Development Economics*, *124*, 199–213.

Bluffstone, R., Dannenberg, A., Martinsson, P., Jha, P., & Bista, R. (2020). Cooperative behavior and common pool resources: Experimental evidence from community forest user groups in Nepal. *World Development*, *129*, Article 104889.

Brañas-Garza, P., Cobo-Reyes, R., & Domínguez, A. (2006). "Si él lo necesita": Gypsy fairness in vallecas. *Experimental Economics*, *9*(3), 253–264.

Burki, S. J., Perry, G., & Dillinger, W. R. (1999). *Beyond the center: decentralizing the state*. World Bank Publications.

Cadsby, C. B., & Maynes, E. (1999). Voluntary provision of threshold public goods with continuous contributions: experimental evidence. *Journal Public Economics*, *71*(1), 53–73.

Camerer, C. (2015). The promise and success of lab-field generalizability in experimental economics: A critical reply to levitt and list. In G. Fréchette, & A. Schotter (Eds.), *Handbook of experimental economic methodology*.

Carlsson, F., Johansson-Stenman, O., & Nam, P. K. (2014). Social preferences are stable over long periods of time. *Journal of Public Economics*, *117*, 104–114.

Corduneanu-Huci, C., Dorsch, M. T., & Maarek, P. (2021). The politics of experimentation: Political competition and randomized controlled trials. *Journal of Comparative Economics*, *49*(1), 1–21.

Croson, R. T., & Marks, M. B. (2000). Step returns in threshold public goods: A meta-and experimental analysis. *Experimental Economics*, *2*(3), 239–259.

d'Adda, G. (2017). Relative social status and conformism: Experimental evidence on local public good contributions. *Economics Letters*, *157*, 31–35.

Dal Bó, P., Foster, A., & Putterman, L. (2010). Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review*, *100*(5), 2205–2229.

Das, S. (2020). (Don't) leave politics out of it: Reflections on public policies, experiments, and interventions. *World Development*, *127*, Article 104792.

del Pilar Moreno-Sánchez, R., & Maldonado, J. H. (2010). Evaluating the role of co-management in improving governance of marine protected areas: An experimental approach in the Colombian caribbean. *Ecological Economics*, *69*(12), 2557–2567.

Fenyvesi, A. (2005). *Hungarian language contact outside hungary: studies on hungarian as a minority language, vol. 20*. John Benjamins Publishing.

Gächter, S., Nosenzo, D., Renner, E., & Sefton, M. (2012). Who makes a good leader? Cooperativeness, optimism, and leading-by-example. *Economic Inquiry*, *50*(4), 953–967.

Gatiso, T. T., & Vollan, B. (2017). Democracy and cooperation in commons management: experimental evidence of representative and direct democracy from community forests in Ethiopia. *Environment and Development Economics*, *22*(2), 110–132.

Gatiso, T. T., Vollan, B., & Nuppenau, E.-A. (2015). Resource scarcity and democratic elections in commons dilemmas: An experiment on forest use in Ethiopia. *Ecological Economics*, *114*, 199–207.

Giné, X., Goldberg, J., Silverman, D., & Yang, D. (2018). Revising commitments: Field evidence on the adjustment of prior choices. *The Economic Journal*, *128*(608), 159–188.

Haas, N., & Khadka, P. B. (2020). If they endorse it, I can't trust it: How outgroup leader endorsements undercut public support for civil war peace settlements. *American Journal of Political Science*.

Harper, K., Steger, T., & Filčák, R. (2009). Environmental justice and roma communities in central and Eastern Europe. *Environmental Policy and Governance*, *19*(4), 251–268.

Isaac, R. M., & Walker, J. M. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *Quarterly Journal of Economics*, *103*(1), 179–199.

Jack, B. K., & Recalde, M. P. (2015). Leadership and the voluntary provision of public goods: Field evidence from bolivia. *Journal of Public Economics*, *122*, 80–93.

Javaid, A., & Falk, T. (2015). Incorporating local institutions in irrigation experiments: evidence from rural communities in Pakistan. *Ecology and Society*, *20*(2).

Kaczan, D., Pfaff, A., Rodriguez, L., & Shapiro-Garza, E. (2017). Increasing the impact of collective incentives in payments for ecosystem services. *Journal of Environmental Economics and Management*, *86*, 48–67.

Kocher, M. G., Martinsson, P., & Visser, M. (2008). Does stake size matter for cooperation and punishment? *Economics Letters*, *99*(3), 508–511.

Kosfeld, M., & Rustagi, D. (2015). Leader punishment and cooperation in groups: Experimental field evidence from commons management in Ethiopia. *American Economic Review*, *105*(2), 747–783.

Kostadinova, T. (2009). Abstain or rebel: Corruption perceptions and voting in east European elections. *Politics & Policy*, *37*(4), 691–714.

Ledyard, J. (1995). Public goods: A survey in experimental research. In *Handbook of experimental economics* (pp. 111–194).

Levati, M. V., & Morone, A. (2013). Voluntary contributions with risky and uncertain marginal returns: The importance of the parameter values. *Journal of Public Economic Theory*, *15*(5), 736–744.

Markussen, T., Putterman, L., & Tyran, J.-R. (2014). Self-organization for collective action: An experimental study of voting on sanction regimes. *Review of Economic Studies*, 301–324.

Martín, J., Branas-Garza, P., Espín, A. M., Gamella, J. F., & Herrmann, B. (2019). The appropriate response of spanish gitanos: Short-run orientation beyond current socio-economic status. *Evolution and Human Behaviour*, *40*(1), 12–22.

Martinsson, P., Villegas-Palacio, C., & Wollbrant, C. (2015). Cooperation and social classes: Evidence from Colombia. *Social Choice and Welfare*, *45*(4), 829–848.

Matsusaka, J. G. (2005). Direct democracy works. *Journal of Economic Perspectives*, *19*(2), 185–206.

Mendizabal, I., Valente, C., Gusmão, A., Alves, C., Gomes, V., Goios, A., et al. (2011). Reconstructing the Indian origin and dispersal of the European Roma: A maternal genetic perspective. *PLoS One*, *6*(1).

Murrell, P. (2008). Institutions and firms in transition economies. In *Handbook of new institutional economics* (pp. 667–699). Springer.

Mutz, D. C., Pemantle, R., & Pham, P. (2019). The perils of balance testing in experimental design: Messy analyses of clean data. *The American Statistician*, *73*(1), 32–42.

Olken, B. A. (2010). Direct democracy and local public goods: Evidence from a field experiment in Indonesia. *American Political Science Review*, *104*(2), 243–267.

Ostrom, E. (1990). *Governing the commons: the evolution of institutions for collective action*. Cambridge University Press.

Pereda, M., Tamarit, I., Antonioni, A., Cuesta, J. A., Hernández, P., & Sánchez, A. (2019). Large scale and information effects on cooperation in public good games. *Scientific Reports*, *9*(1), 1–10.

Péteri, G. (2018). Decentralization in Eastern Europe: grab the moment. Website of the MTA–DE Public Service Research Group.

Reindl, I., Hoffmann, R., & Kittel, B. (2019). Let the others do the job: Comparing public good contribution behavior in the lab and in the field. *Journal of Behavioral and Experimental Economics*, *81*, 73–83.

Rodríguez-Pose, A., & Krøijer, A. (2009). Fiscal decentralization and economic growth in central and eastern europe. *Growth and Change*, *40*(3), 387–417.

Schram, A., Offerman, T., & Sonnemans, J. (2008). Explaining the comparative statics in step-level public good games. In *Handbook of experimental economics results, vol. 1* (pp. 817–824). Elsevier.

Somanathan, E., Prabhakar, R., & Mehta, B. S. (2009). Decentralization for cost-effective conservation. *Proceedings of the National Academy of Sciences*, *106*(11), 4143–4147.

Tognetti, A., Berticat, C., Raymond, M., & Faurie, C. (2012). Sexual selection of human cooperative behaviour: An experimental study in rural Senegal. *PLoS One*, *7*(9).

Viceisza, A. C. (2016). Creating a lab in the field: Economics experiments for policymaking. *Journal of Economic Surveys*, *30*(5), 835–854.

Zarychta, A. (2020). Making social services work better for the poor: Evidence from a natural experiment with health sector decentralization in honduras. *World Development*, *133*, Article 104996.