Evolution and Human Behavior 43 (2022) 53–70

Contents lists available at ScienceDirect

Evolution and Human Behavior

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

Institutional quality shapes cooperation with out-group strangers

Marco Fabbri *

Department of Economics and Business, Aarhus University, Denmark & Amsterdam Center for Law and Economics, Netherlands

ARTICLE INFO

Keywords:
Cultural Evolution
In-group Favoritism
Lab-in-the-field Experiment
Material Security Hypothesis
Multilevel Public Goods Game
Parochialism

ABSTRACT

Humans display a puzzling cross-population variation in the ability to cooperate with out-group members. One hypothesis is that impartial institutions substituting kith and kin as risk-buffering providers would favor the expansion of cooperative networks. Here I propose a research design that overcomes the endogeneity between institutions and preferences, making it possible to isolate the causal effects of institutional quality on out-group cooperation. I study a land tenure reform implemented as a randomized control-trial in hundreds of Beninese villages. The reform reduces the village community's discretion in regulating members' access to land by granting formal legal protection to individual rights-holders. Using a lab-in-the-field incentivized experiment ($N = 576$), I show that the reform significantly increases participants' cooperation with anonymous strangers from other villages. The results illustrate how humans' investments in in-group and out-group relationships are sensitive to cost-benefit evaluations, and emphasize that the institutional environment is a key driver of large-scale human cooperation.

1. Introduction

The ability to engage and sustain cooperation in large networks of unrelated strangers is a feature that makes humans unique within the animal world (Bowles & Gintis, 2013; Maynard Smith & Szathmary, 1999; Seabright, 2010). Nevertheless, there is abundant evidence to show that individuals tend to display significantly higher levels of cooperation when interacting with members of their own reference group, regardless of whether group membership is defined by kinship (Schulz et al., 2019), ethnicity (Bernhard, Fischbacher, & Fehr, 2006; Bowles & Gintis, 2004), active role or organization affiliation (Goette, Huffman, & Meier, 2006), religion (Lang et al., 2019), or even minimal group labels in anonymised laboratory experiments (Chen & Li, 2009). However, limiting interactions to kith and kin might prevent the development of large networks which would enhance specialization, economies of scale, and the realization of possible gains from trade, as well as fostering ethnic discrimination and social segregation (Arrunada, 2012; Habyarimana et al., 2007). Therefore, understanding the determinants of humans' tendency to rely on cues of social closeness in guiding cooperative behavior, which I will refer to as "in-group favoritism", and the factors that favor the development of large-scale cooperation, is a key task for increasing the well-being of societies.

A puzzling characteristic of in-group favoritism is found in its large cross-population variation (Fiedler et al., 2018; J. Henrich & Muthukrishna, 2021). Several hypotheses have been advanced to account for differences in levels of cooperation with out-group individuals across societies, including the effects of exposure to markets and globalization (Buchan et al., 2009; J. Henrich, Ensminger, et al., 2010), environmental conditions (de Vliert and Evert, 2020), inter-group conflicts (Handley & Mathew, 2020), pathogens stress (Fincher & Thornhill, 2012), or the influence of religious beliefs on kinship networks (J. Henrich, 2017; Schulz et al., 2019). In ad- dition to these hypotheses, a recent wave of research both in anthropology and in economics has focused on the effects that the quality of institutions has on enlarging the scope of cooperation with outgroups. One view proposed in evolutionary anthropology is that strong norms of intra-group cooperation, paralleled by distrust toward outgroups, represent a risk buffering strategy against existential threats (Hruschka & Henrich, 2013). Therefore, according to this view, the creation of impartial institutions capable of mitigating these threats would scale up cooperation: well-functioning institutions substitute kith and kin as social insurance providers, thus relaxing the need for an individual to maintain strong in-group relationships and opening up the possibility of expanding one's own network (Powers, van Schaik, & Lehmann, 2016). Scholars in economics have also suggested that well-functioning institutions, ones that impartially apply enforcement *erga omnes* without limiting it to local situations, would enhance values of generalized morality and expand the scope of cooperation (Tabellini, 2008). In particular, most

* Corresponding author.
E-mail address: marco.fabbri@econ.au.dk.

https://doi.org/10.1016/j.evolhumbehav.2021.11.003
Received 19 April 2021; Received in revised form 18 October 2021; Accepted 2 November 2021
Available online 22 November 2021

studies of this literature focus on the effects that institutions have on people's values and on the coevolutionary process of these values and of the institutions that a population choose (Acemoglu & Robinson, 2019; Bisin & Verdier, 2001; Lowes et al., 2017; Tabellini, 2010).

This paper contributes to this body of research studying the institutional determinants of in-group favoritism. I investigate whether the introduction of more impartial institutions affects the willingness to cooperate with out-group strangers. I do so by studying a land tenure reform in rural Benin that formalizes property rights over land, making them enforceable in state courts and replacing the previous system where access to land was informally granted and secured through relationships with the local community members. I test whether the replacement of the socially-determined informal norms of access to land with a system of formal and impartially enforced land tenure institutions produce an increase in individuals' willingness to cooperate with out-groups. One feature that characterizes my approach is the focus on isolating the unidirectional causal effects of impartial institutions in expanding the scope of cooperation, focusing on short-term changes (i.e. within the time span of one generation). Moreover, while my study is not designed to evaluate the hypothesis mentioned above regarding other possible determinants of in-group favoritism, throughout the analysis I estimate whether these factors interacted with, or mediated the effects of, the institutional reform I am studying. I will come back to this point in the discussion of the mechanisms presented in Section 3.

Notwithstanding the abundant literature that examines the causes of in-group favoritism (Balliet, Junhui, & De Dreu, 2014), relatively few attempts have been made to investigate empirically the relationship between institutional quality and cooperation with out-groups, delivering mixed results. Evidence from cross-population analysis based on observational data shows that institutional quality and material security are associated with lower levels of in-group favoritism (i.e. the Material Security Hypothesis; Hruschka & Henrich, 2013; Muthukrishna, Henrich, & Slingerland, 2021). Similarly, evidence from laboratory and lab-in-the-field experiments comparing behavioral traits across small-scale societies show that the willingness to cooperate with out-group members is related to an individual's perception of risks and their living conditions (Hruschka, Efferson, et al., 2014; Kaplan et al., 2012). However, recent evidence suggests that institutional quality and market integration do not correlate with parochial trust in the context of an online experiment across 17 societies (Romano et al., 2017). Moreover, results from non-anonymised lab-in-the-field games show that those Bolivian villagers who have comparatively higher socio-economic status and access to markets are less cooperative with out-group strangers (Pisor & Gurven, 2016).

Despite the important insights conveyed by this wave of research, empirical investigations into the institutional determinants of parochialism face a major challenge which consists of isolating the direction of the causal effects that institutions have on in-group favoritism. While cross-population studies can be useful for looking into and sorting out different hypotheses, establishing causal relationships requires the observation of institutional changes that are orthogonal to individuals' preferences for in-group cooperation. This is problematic because social institutions are endogenous constructs that reflect the preferences of the institution builders and co-evolve with those preferences (Greif & Laitin, 2004). Furthermore, comparing different populations cannot account for the self-selection of individuals or of their ancestors, as human migration is not random and genetic legacy or culturally-transmitted behavioral traits can persist for generations after formal institutions have disappeared (Boyd & Richerson, 2009; MacDonald, 2019). A similar problem is faced by the approach that treats institutional changes, such as historical transitions of state regimes, legal reforms, or regulatory modifications, as "natural experiments" which can be used to estimate their effects on people's preferences. Assuming that those institutional modifications are exogenous to individuals' culture does not dispel endogeneity concerns, since it is often impossible to establish whether the replacement of existing institutions is the cause or the consequence of the mutated preferences of the institution's builders (Alesina & Giuliano, 2015). Moreover, although laboratory experiments that study the reaction of subjects randomly exposed to different institutional environments would solve the identification problem, the approach faces serious limitations because of the impossibility of studying the long-term effects of exposure to different institutional settings and the contrived nature of the stylized institutions manipulated in laboratory games (J. Henrich, Heine, & Norenzayan, 2010).

The research design implemented in this article overcomes the identification challenge linked to the endogeneity that characterizes real-world institutions. The institutional shock I study is a unique case of large-scale land rights reform that was implemented as a randomized control-trial (RCT). In Benin, as in the vast majority of the African continent, customary tenure characterized by collective property and informal possession largely predominates in rural areas. Access to land is regulated by customary land rights, a complex set of unwritten rules and procedures tightly related to the socio-political dynamics of the village network. These socially-determined land rights grant to the village community and its authorities considerable discretion in allocating, limiting, and revoking land-use rights to individual members (Deininger & Feder, 2009).

Attempting to improve access to land, security of tenure, and the development of a land market, between 2009 and 2011 the Beninese government, with the support of the Millennium Challenge Corporation, implemented an approach for systematic identification and registration of customary rights to parcels of agricultural land, the "Plan Foncier Rural" (PFR). PFR consists of socio-land surveys at the village level to identify rights-holders, define rights, and demarcate land parcel boundaries. Registered rights are stored in publicly-available registries and recorded in certificates that are distributed to rights-holders.[1] The registration awards to rights-holders the formal presumption of ownership recognized by courts, making it possible to sell or to use registered parcels as collateral (Goldstein et al., 2018). Appendix A provides a detailed summary of the legal and operational characteristics of PFR. Notably, the reform has produced a major change in the degree of security of tenure for individuals. According to customary law, a villager's access to land and socially-determined rights of use were subject to the discretion of the village community, whereas following the registration the rights-holder is granted legal protections by the state-enforced formal legal system (Fabbri & Dari-Mattiacci, 2021). To confirm that the increase in legal protection determined by the reform was followed also by an increase of perceived tenure security, I conducted a survey among villagers. This showed that, for instance, over 90% of respondents believe PFR registration increases tenure security and that the legal enforcement provided by official state courts is considered to be the ultimate source of conflict adjudication. Appendix C discusses the survey details.

The Benin PFR is the first case of land tenure reform implemented as a randomized control trial on a large scale. First, 600 eligible villages willing to implement the reform were identified. Eligibility was determined by village characteristics such as being located in rural areas, possessing comparable income levels, and belonging to provinces where other projects related to land rights had not been, or were not in the process of being, implemented. Second, a sub-sample of 294 villages was selected via public lottery, and PFR was actually implemented ("treated"). Non-selected villages ("control") did not receive any intervention and, as of today, continue to have customary land rights. The left panel of Fig. 1 summarizes the implementation procedures, and the right panel indicates the resulting distribution of treated and control villages across the country. The random sorting through a lottery mechanism of eligible villages in the two groups cancels out confounding factors. To

[1] In 2013 the automatic distribution of certificates was replaced by a new procedure for requesting a titling document while leaving the validity of registered rights unaffected.

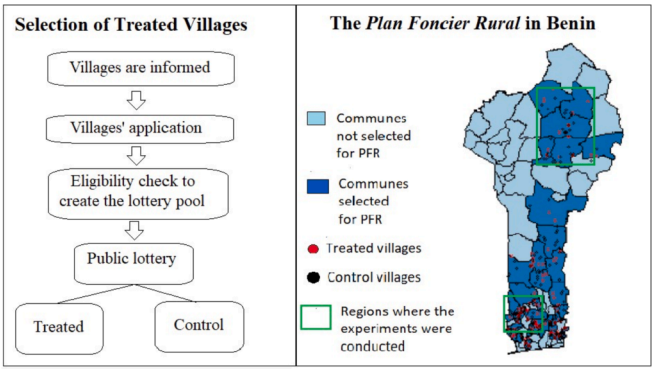

Fig. 1. *Left Panel.* The lottery mechanism characterizing the Beninese Plan Foncier Rural. Following a campaign organized by the government to inform the village authorities about the details of the reform, the interested villages could submit an application for having it imple- mented. Once the applications were received, they went through an eligibility check that took into account the following criteria: poverty index, potential for commercial activities, regional market integration, local interest in promoting gender equality, infrastructure for economic activities, adherence to the PFR application procedure, the incidence of land conflicts, and the production of main crops. Moreover, villages had to be located in rural areas and in communes (i.e. provinces) where other projects related to land rights had not been, or were not in the process of being, implemented. *Right Panel.* The resulting distribution of treated and control villages across the country, indicating the two regions where the fieldwork took place.

isolate the causal relationship between securing property rights and expanding cooperative networks, I compare the level of cooperation with out-groups elicited by using an incentivized experiment involving 576 participants equally distributed between 32 villages included in the PFR lottery pool.

To measure out-group levels of cooperation, my research team conducted a lab-in-the-field economic experiment that replicates (with some minor variations) the one first proposed by Buchan et al., 2009. The experimental design resembles that of a multilevel public goods experiment, except that subjects do not make decisions directly affecting only those in their concurrent groups in the session. The details of the experimental protocol can be found in the Appendix B.1. In this game, participants have to decide how to allocate 10 coins worth XOF 100 each (roughly $0.17), which they have received from the experimenter, between three accounts: an "Individual" account, a "Village" account, and a "Country" account. Coins allocated to the Individual account are kept by the participant. Coins allocated to the Village account will be added to the coins allocated to the Village account of two other participants belonging to the same village. The experimenter then increases the total amount of coins in the Village account by 50% and divides it equally between the three participants, thus entailing a half-token loss for the contributor and yielding a return of 0.5 tokens for the other members of the Village group. Coins allocated to the Country account by each participant are added to the coins that two other participants from the same village contributed to the Country account, and to the coins allocated to the Country account by six participants belonging to different villages in Benin. The total amount of coins in the Country account is then doubled by the experimenter before being equally divided among the nine participants. Therefore, each coin allocated to the Country account entails a 0.78 token-loss for the individual and yields a return of 0.22 tokens to each of the other eight group members. The final earnings of each participant are the sum of the coins allocated to the Individual account, their share of the coins coming from the Village account, and their share of the coins coming from the Country account.

This incentives structure mirrors that of a multilevel public goods dilemma. A purely self-interested individual would allocate all their tokens to the Individual account, since the returns from coins allocated to both the Village account (0.5) and the Country account (0.22) are lower. However, it is well-known that in public goods experiments people deviate from the selfish strategy of non-contribution and allocate positive amounts to the common account (Zelmer, 2003). Moreover, subjects that would like to contribute to the public good face an additional trade-off when deciding for the Village or Country account. On the one hand, the Village account has a higher return per token contributed (0.5) than the Country account (0.22). On the other hand, if each group member contributes their full endowment to the Country account, the final individual payoff would be larger (20 coins) than the one obtained by full contribution to the Village account (15 coins).

The game design aims to reproduce the relations between same-village members and non- locals in a laboratory setting. The multilevel public goods structure reflects the idea that impersonal property institutions do not exclude an individual from engaging in local interactions. Instead, formal legal enforcement reduces some of the risks associated with expanding cooperative networks to non-locals, such as having individual land expropriated by the village community because of fellow villagers' antipathy towards business with strangers. For the same reason, the experiment proposes the original design feature proposed by Buchan et al., 2009, according to which a participant's contribution to the Country account also benefits two fellow villagers. While this design feature might to some extent blur the distinction between in-group and out-group members, it captures the frequently nested nature of cooperative networks operating at different levels. In the analysis, I consider a contribution to the Village account as reflecting in-group interests, while contribution to the Country account and the interaction with individuals from different villages are considered as mirroring out-group interests. The analysis predicts how contributions to the Country account are affected by the creation of more impersonal property institutions.

2. Materials and methods

A pre-analysis plan, describing the main hypotheses to be tested, the experimental design, data collection procedures, and the empirical strategy, was posted on the AEA RCT registry before the data collection took place (see RCT ID AEARCTR-XXXX). There were no deviations from the pre-registered plan.

2.1. Participants

The data collection took place in the provinces of Mono and Couffou (in the South of the country) and Alibori and Borgou (in the North) between January and March 2020. With the support of local research assistants, I carried on the fieldwork activity collecting lab-in-the-field experimental and survey data on 32 villages (16 treated and 16 control) randomly selected among those included in the lottery pool in these provinces. The research assistants had no prior knowledge of the hypothesis to be tested. We ran a session with 18 participants in each village visited, for a total of 576 participants who took part to the experiment. During one session, one participant left the experiment because of feeling sick, so we actually have data relative to 575 individual choices.

The fieldwork procedure was as follows. The day before the experiment, a research assistant met the village chief to inform him or her that the following day a team of researchers would come to the village. The village chief was requested to inform the population and gather at the designed time as many volunteers as possible for participating in the study. Participants must be resident in the village, of age, and only one member per household was allowed to participate. On the day of the experiment, the research team randomly selected nine male and nine female participants among the volunteers who showed up and met these criteria. Since the fieldwork was completed during the dry season when most of the agricultural work is suspended, there were always more than 18 volunteers willing to participate. Non-selected participants were paid a show-up fee of XOF 500 (approximately $0.8) and requested to leave.

2.2. Materials and procedure

The project was reviewed and received approval by the Research Ethic Committee "Parc de Salut Mar Barcelona" (ref. 2018/8015/I). All experiments were performed in accordance with relevant guidelines and regulations. Informed consent was obtained from all participants.

The experimental sessions took place in a public space (usually a school or the village meeting room). The instructions were read aloud by the experimenter in a common room and the decisions were privately taken by each subject in a separate decision room. Before entering the decision room, the experimenter asked each participant in private a set of questions to verify the comprehension of the game (a copy of the instructions, the comprehensions checks, and visual support used during the experiment are included in Appendix D). If the participant could not answer a question, the experimenter repeated the instructions until the participant was able to answer correctly.

Participants were asked to make the contribution decisions in the multilevel public goods game described in Section 1. When called to make decisions during the experimental session, the participant entered the decision room alone and physically divided the ten coins between the different envelopes, which were marked with a code. This procedure made the contribution decisions of the participants blind to the experimenter on site.

Participants took also part in other pre-registered experimental games whose outcomes were used in the analysis. First, they made a contribution decision in a standard linear public goods game played with fellow villagers (the game details are specified in Appendix B.1). As specified in the pre-registered analysis plan, I follow Buchan et al. (2009) and use the contribution in this game as a proxy that allows me to control for baseline levels of individual cooperation. Second, participants took part to a Dictator game framed as a donation to an unspecified orphanage outside the village. Each participant was given 10 tokens worth XOF 100 each and had to choose how many coins to keep for herself and how many to donate to the orphanage. Third, they completed an incentivized risk elicitation task. Each participant was given six coins. The participant could invest part or all of the coins in a bet in which with 50% probability the amount invested will be doubled, or lost. The coins not invested were added to the participant's final payments. Finally, they also completed a socio-demographic survey, reporting the following pre-specified information: age, gender, religion, marital status, number of family members, participation in household finance management, education, literacy, village of birth, years of residence in the village, and measures of income.

Before receiving the final payment, participants took also part to other fieldwork activities not related to this project. The public goods games described in this paper were the first activity conducted during the fieldwork in each session. Until the public goods games had terminated, participants were not informed that they would take part to the additional activities. Each field-work session lasted approximately three hours. Participants received as a final payment on average XOF 2800 ($4,5), roughly the equivalent to the wage earned in one and a half days of work for the median subject in our sample.

2.3. Data analyses

The data and replication files were submitted together with the article and are available at the publisher's website as supplementary material. The data analyses were performed using the statistical software STATA (version 16). As a preliminary step in the analysis, I check the internal validity of the research design, which is based on the RCT implementation of the Plan Foncier Rural reform. For the validity of the results presented in the next section to be upheld, two caveats are in order. First, I need to verify that the implementation of the PFR reform across the 575 villages included in the lottery pool was effectively random. Similarly, I need to check that the selection of the sample of participants drawn from the 32 villages where the data collection for this study took place was not affected by accidental imbalances. Second, for the identification strategy to work, I must ensure that there is no selection into treatment following the implementation of the reform (for instance, because of migration from control to treated villages after the reform implementation). Regarding the latter point, in the next section I report checks performed to verify absence of selection into treatment. Regarding the former point, pre-treatment data on observables and cooperation relative to participants in the study are not available, thus making it impossible to perform a pre-post treatment balance comparison. To mitigate concerns regarding accidental imbalances from randomization, as a preliminary step I test for differences of participants' observables across treatments in my sample.

The analysis continues by testing the main hypothesis. In the baseline empirical specification, the dependent variable is the amount of coins contributed to the Country account. The regressors include the treatment dummy, the proxy for baseline individual cooperation, the socio-demographics listed in subsection 2.2, and dummies for geographical region, certificates distributed before the 2013 change in title-requesting procedure, and villages that have expanded or modified the original intervention. I always report the results of both one specification controlling for education and literacy, and a separate specification which excludes these regressors, since previous research has shown that investments in education can be endogenous to being awarded private property (Galiani & Schargrodsky, 2010). In all regression models the standard errors are clustered at the village level. I report the results of three regression models. First, I present the results of OLS regressions, which provides a straightforward interpretation of the treatment effect. Second, I show the results of a left-censored Tobit regression in order to take into account the accumulation of observations at zero contribution. Third, I report the results of a double-hurdle regression for continuous data (Engel & Moffatt, 2014). This model relaxes the over-restrictive implicit assumption of the Tobit model that there is only one type of zero-contribution observation, always deriving from subjects' circumstances and treatment. The double-hurdle model embodies the idea that individuals' decision to contribute to the public goods is the result of two processes: the first hurdle, determining whether the individual is a "zero-contribution" type, and the second hurdle, determining the extent of contribution given that the individual is not a zero-contribution type. Therefore, a key feature of the model is that it distinguishes between two types of zero contribution observations: on the one hand, if the individual is a zero-contribution type, the contribution to the public goods will always be zero, no matter what are the circumstances experienced by the participant at the time of the decision; on the other hand, the individual might not be a zero-contribution type, but the current circumstances and treatment conditions might result in a zero contribution. This sort of zero is usually classified as "censored zero after" (Tobin, 1958). I estimate the outer hurdle (i.e. whether a subject is a zero-contribution type) by using the proxy for baseline individual cooperation elicited in the separate standard public goods game and time-invariant individual characteristics – age and gender. I exclude the proxy for baseline cooperativeness from the estimation of the inner hurdle. In the robustness checks, I show that the inclusion or exclusion of baseline cooperativeness in the outer or inner hurdles does not affect the qualitative results.

The analyses continue by testing the robustness of the main results. First, I perform several robustness checks of the regression models described above by controlling for additional individual and village-level characteristics. Second, I employ a Lasso post-double-selection methodology for appropriately selecting the controls to be included in the regression (Belloni, Chernozhukov, & Hansen, 2014). This method has been proved useful to improve the robustness of causal inference when accidental imbalances in the sample occur (Chernozhukov et al., 2018). Third, according to the 2009-2011 PFR formulation, only parcels of land within the administrative village boundaries were subject to the intervention. However, in three villages of the treated group, the village authorities reported having further extended the original PFR

intervention after 2011 by registering some additional land parcels lying outside the official village boundaries. Moreover, in one village of the control group that shares borders with a treated village, half of the participants reported having land parcels located within the boundaries of the adjacent village, and so that those parcels were included in the PFR intervention. Therefore, I verify whether the exclusion of these villages from the sample affects the results

As a final step in the analysis, I try to shed light on possible mechanisms that might have determined the observed results. First, I check whether differences in out-group cooperation are driven by participants in control villages displaying a lower trust in government authorities as a reaction to not having being selected by the PFR lottery. Second, I have discussed the role possibly played by conflicts in shaping inter and intra-group cooperation. The reform might have affected the conflict rate experienced by participants, and that could have in turn determined the observed change in out-group cooperation. Therefore, I test whether changes in the rate of land-related conflicts, following the reform implementation, mediated the results. Third, the introduction of formalized land rights might have created the possibility for villagers to use land as collateral and so increased participants' access to the credit market. This, in turn, could have facilitated market transactions with unknown strangers and it could have determined the increase in out-group cooperation (Devijlder & Schoors, 2019). I test whether controlling for different measures of market integration, such as the share of calories purchased in markets and distance from markets, capture the observed increase in out-group contribution. Moreover, I test whether participants in treated villages benefited of an improved access to credit. Fourth, I test for possible changes in participants' generalized altruism. To do so, participants take part in a Dictator game framed as a donation to a charity outside the village. Finally, I analyze the data on contribution to the public goods shared with in-group members. I look at both participants' contribution decisions in the standard linear public goods played with fellow villagers, and the contribution to the Village account in the multilevel public goods game. The objective of this analysis is to compare the cooperation decision with in-groups and with out-groups to shed some light on how experiencing the reform affected participants' parochialism in cooperation.[2]

3. Results

3.1. Preliminary checks

I start by reporting results of the analyses performed as preliminary checks for the internal validity of the research design. Regarding the selection of 300 villages to have the reform implemented from the entire RCT pool of 576 villages, an early impact evaluation uses pre-treatment and post-treatment data to show that the randomization was successful (Goldstein et al., 2015). The sample of participants taking part to this study was drawn from 32 villages randomly selected from the whole list of villages included in the PFR that are located in two provinces in the north of the country (Mono and Couffou) and two provinces in the south (Alibori and Borgou). Table 1 shows that participants in the sample are in general well-balanced across treatment branches, with the exception of some differences in average age, likelihood to be married, households with running water, and possession of a bank account.

Regarding participants' self-selecting into treated villages through migration, as shown by looking at survey data discussed in Appendix C, migrating out of the village of origin is rare for subjects in my sample and migration flows are similar across treatment branches. Only 35 out of 575 participants were not already resident in the village when the PFR reform was implemented, 20 in treated villages and 15 in control. The difference in not statistically significant (χ^2 test, p>10%). The majority of these migrations were reported by female participants, and the stated reason was marriage in over 90% of the cases. Moreover, the regression analysis below controls for village of origin and years spent in the village where the data were collected.

Table 1
Balance of observables across treatment groups (t test two-sided for continuous variable and chi-square test for dummy variables).

	Treated	Control	Difference
	(n = 288)	(n = 287)	(p-value)
male	0.49	0.51	0.67
age	40.1	36.8	0.01
muslim	0.45	0.39	0.09
vodoun	0.19	0.21	0.40
other religion	0.03	0.02	0.43
married	0.89	0.83	0.02
education	1.35	1.20	0.24
literacy	0.40	0.35	0.27
nr. households' components	9.8	9.5	0.57
manage household's finance	0.96	0.95	0.52
born in the village	0.69	0.73	0.23
fraction of lifetime in the village	0.80	0.83	0.12
weekly logincome (XOF)	8.28	8.37	0.59
land (hect)	20.84	20.73	0.97
concrete floor	0.64	0.57	0.11
electricity	0.37	0.36	0.86
water	0.26	0.18	0.01
radio-TV	0.63	0.59	0.35
car	0.09	0.06	0.21
moto	0.78	0.77	0.55
bank account or credit card	0.33	0.25	0.04
self-reported social-rank	4.45	4.32	0.39

3.2. Main result

The analysis begins by plotting in Fig. 2 the average amount of coins contributed to the Country account by treated and control villagers. Participants in treated villages contributed on average 3.4 coins to the Country account while those in the control group contributed 2.9 coins. The difference between the two groups is statistically significant at the conventional level (t-test two-sided, p=0.036).

Table 2 reports the results of the regression analysis estimating the basic specifications described in Section 2. The amount of coins contributed to the Country account is the dependent variable. Models 1 and 2 report OLS results. In Model 1, the coefficient of the treatment dummy is positive and statistically significant at the conventional level,

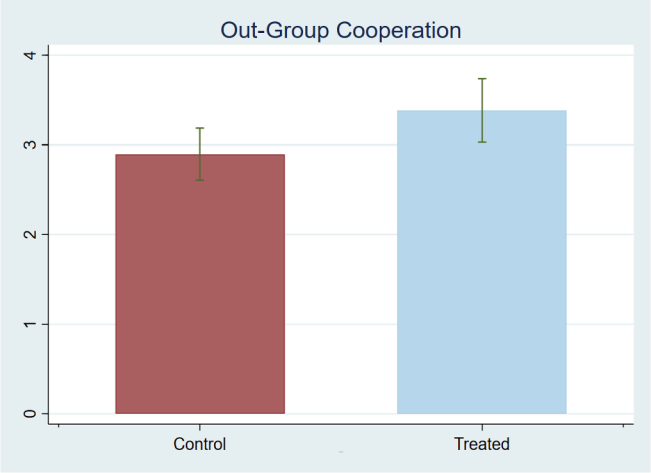

Fig. 2. Coins contributed to the Country account. Vertical bars display 95% confidence intervals.

[2] I also tested possible heterogeneous treatment effects on out-group cooperation in terms of gender and income levels. Results (available upon request) show uniform treatment effects across gender and income levels.

Table 2
Out-group cooperation.

	Model 1 OLS	Model 2 OLS	Model 3 Tobit	Model 4 Tobit	Model 5 Hurdle	Model 6 Hurdle
treated	1.018*	0.996*	1.103*	1.086*	1.339*	1.385*
	(0.485)	(0.487)	(0.546)	(0.550)	(0.669)	(0.688)
coop-ingroup	0.308***	0.308***	0.344***	0.344***		
	(0.088)	(0.087)	(0.097)	(0.096)		
education		−0.089		−0.097	−0.128	
		(0.148)		(0.175)	(0.176)	
literate		0.653		0.626	0.886	
		(0.520)		(0.579)	(0.655)	
Constant	−0.459	−0.957	−1.524	−1.951	2.940***	2.865***
	(1.234)	(1.387)	(1.495)	(1.633)	(0.280)	(0.0.258)
var(e.coop2-foreign)			9.403***	9.350***		
			(1.183)	(1.183)		
N.obs.	575	575	575	575	575	575

Notes: Dependent variable: Contribution to the Country account in decision 2. Models 1 and 2: OLS regression. Models 3 and 4: Left Censored Tobit regression. Models 5 and 6: Double-hurdle regression. Standard errors robust for clustering at the village level. Models 5 and 6 implement bootstrapping with 50 repetitions. All models include as controls: age, gender, religion, marital status, number of family members, participation in household finance management, village of birth, years of residence in the village, self-reported weekly income, two dummies for villages where the PFR was extended after 2011, a dummy for participants who received a land-rights certificate, a dummy for villages in the South. Compared to Models 1, 3, and 5, Models 2,4, and 6 additionally controls for literacy and education levels. In Models 1-4 the variable *coop-ingroup* reports the coins contributed in a separate standard linear public goods game and is used as regressor. In Models 5 and 6 the explanatory variables for the outer hurdle are *coop-ingroup, age, and gender*. Symbols ***, **, and * indicate significance at the 0,1%, 1% and 5% level, respectively.

suggesting that participants in treated villages contribute significantly more to the Country account compared to those in the control group. The coefficient of the variable *coop-ingroup* indicating the amount of coins contributed in the standard public goods game is also positive and significant. This result should not come as a surprise, since it actually confirms that subjects are consistent across decision settings. In model 2, in which education and literacy are added as controls, the qualitative results remain unchanged compared to the ones of the main specification, and the point estimates are very similar. Models 3 and 4 report the results of a left-censored Tobit estimation. The coefficient of the treatment dummy remain positive and significant at the conventional level, confirming the results.[3]

Models 5 and 6 report the estimated results of a double-hurdle model which uses the method of maximum simulated likelihood (Engel & Moffatt, 2014). I cluster standard errors at the village level and implement bootstrapping with 50 repetitions. The results of the specification presented here model the first hurdle (i.e. whether or not being a zero-contribution type) as being affected by baseline individual cooperation, age, and gender. The second hurdle (i.e. how many coins contributing, providing one is not a zero-contribution type) does not include baseline cooperation. The coefficient of the treatment dummy is again positive and significant at conventional level. The interpretation changes slightly from the Tobit and OLS: the coefficient estimates the effect of being in a village where the reform was implemented on the contribution to the Country account, on condition of passing the first hurdle, that is, on not being a zero-contribution type. In Table A1 in Appendix C, I re-estimate the model specifications in Table 2 by a) including the proxy for baseline cooperation in the inner hurdle instead of the outer one; and b) excluding baseline individual cooperation from both the inner and outer hurdles. In both cases, the point estimates remain very similar and the qualitative results are the same or better, when compared to the model specifications presented in the main text.

The analysis continues by checking whether the results are sensitive to the inclusion of additional controls.[4] Models 1 and 2 of Table A2 in Appendix C include 14 proxy measures for individual wealth. The result remains qualitatively the same. Models 3 and 4 include additional village-level controls such as population, distance to public healthcare facilities, and whether the village has a school. Results remain qualitatively the same and quantitatively similar to those in the baseline specification.I then re-estimate the model specification presented in Table 2 employing the Lasso post-double-selection approach (Belloni et al., 2014). Table A3 in Appendix Appendix C reports the results. The coefficient of the treatment dummy is confirmed to be significant at the conventional level, and the point estimate to be similar compared to the pre-specified regression, both in the base model specification (model 1) and when we exclude four villages that have extended the PFR plan after the 2010–2011 intervention (model 2). Finally, I verify how estimation results change when four villages that have partially modified the original PFR are excluded from the sample. Fig. A1 in Appendix C replicates Fig. 2 in the main text excluding these villages. Table A4 in Appendix C re-estimates the model specifications presented in Table 2 in the main text excluding these villages from the sample. Results remain qualitatively unchanged.

3.3. Possible mechanisms and heterogeneous effects

One possible explanation for the difference in contribution observed across treatments could be that control participants display disappointment, and associated reduction in trust of government authorities who organized the selection process, for not having been selected by the PFR lottery. In the post-experimental survey, participants were asked to report their appreciation of the government's action and their trust in the impartiality of the formal judiciary vis-à-vis the customary dispute

[3] Differently than the OLS coefficient of the variable *treated*, which reports the effect of being in a PFR village on contribution to the Country account, the Tobit coefficient of the treatment dummy estimates a combination of the change in contribution of those who contributed, weighted by the probability of contributing; and the change in the probability of contributing, weighted by the expected value of contribution.

[4] For the analyses described in what follows, I report the results only of OLS (which provide a straightforward interpretation of the estimated coefficient) and double-hurdle models (since it is an extension of the Tobit model that better suits the data-generating process analyzed). Results of left-censored Tobit regressions are quantitatively and qualitatively very similar and available upon request.

resolution authorities (details are reported in Appendix C). In both cases, no significant difference between treatment branches is found in the rate of replies. Similarly, the contribution to non-local public goods for participants in the experiment does not differ for those who display high and low levels of appreciation for the government, suggesting that this factor is unlikely to have determined the results.

I then verify the possibility that changes in land-related conflicts following the reform mediate the observed increase in out-group cooperation. To do so, I replicate the specifications pre- sented in Table 2 additionally controlling for whether participants have experienced land-related conflicts in the years following the reform implementation. Results are reported in Table A5 in Appendix C. The point estimates and qualitative results remain unaffected by the introduction of this control (with the exception of OLS model specification 2, where the coefficient of the treatment dummy becomes marginally significant at the 5,3% level). This findings suggest that conflicts do not drive the results.

I test whether participation in market activities and access to credit determined the observed increase in out-group cooperation. First, Table A6 in Appendix C replicates the regression analysis presented in the main text by additionally including two proxies for participants' level of market integration – the share of calories purchased in the market and the distance from the nearest market. Results remain qualitatively unchanged. Additionally, Table A7 in Appendix C reports results of a test for differences across treatment of participants' requests for credit to financial institutions or to friends and family members. The results, in line with findings reported by an earlier study (Goldstein et al., 2018), show that the reform did not significantly affect participants' access to credit, thus suggesting that this channel cannot be responsible for the observed results.

I also test whether the reform determined a more general change in generalized altruism that could have contributed to determine the observed increase in the contribution of non-local public goods. To elicit generalized altruism, I asked participants to play a Dictator game framed as a donation to a charity outside the village. The difference between the coins donated in treated and control villages is not statistically significant. If anything, on average participants in control villages donate slightly more coins.

I then turn to the analyses of data relative to contribution decisions to public goods shared with in-groups. As a first step, I compare contributions to the common account in the separate standard linear public goods game played with fellow villagers. Fig. A2 in Appendix C shows contributions to the common account across treatments. The average contribution to the common account is somewhat lower for participants in treated villages compared to control ones (5.3 vs. 5.6). However, a two-sided t-test suggests that the difference is not statistically significant (p=0.15). Regression analyses confirm this result. Table A8 in Appendix C replicates the model specifications presented in Table 2 using as the dependent variable the amount of coins contributed to the common account in the separate standard linear public goods game. The coefficient of the treatment dummy is small and in no case statistically different than zero. As a second step, I compare the contribution to the Village account in decision 2 across treatment branches. Here again the average contribution to the village public good is slightly higher for control participants (3.11) than for treated ones (2.84), although a t-test shows that the difference is not statistically significant (p=0.19). Regression analyses reported in Table A9 in Appendix C replicate the model specifications presented in Table 2 but for using as a dependent variable the contribution to the local public good in decision 2 and confirms the result.

As a final step, I compare the contribution to the Village account relative to the Country account in the multilevel public goods game. The average contribution of participants in control villages to the Village account is slightly larger than the contribution to the Country account (the difference is 0.22). The opposite is observed in treated villages (-0.56). The difference across treatments is statistically significant at the conventional level (p=.026). I investigate further this result using a regression analysis. I regress the difference between Village and Country account contributions on the treatment dummy and the set of covariates used in the main specification. The results of OLS and Tobit models, reported in Table A10 in Appendix C, provide mixed evidence that treated participants contribute significantly more to the Country account relative to the Village account: the coefficient of the treatment dummy is negative in all model specifications, albeit only marginally significant or not significant according to the specification. The results are confirmed if I exclude from the sample 27 subjects who contributed nothing to the public goods in the two decisions. Taken together, these results suggest that the reform increased individuals' willingness to cooperate with out-groups without significantly affecting cooperation with members of the local community, thus producing an expansion of the scope of cooperation rather than a substitution between in-groups and out-groups. However, with respect to the latter point, a cautionary notice is due since the observed average contributions to the local public goods is lower in treated villages. Thus, it is possible that the experiment does not have sufficient power to detect a small but systematic reduction in local cooperation. I will discuss the implication of these findings in the next section.

4. Discussion

We are still far from a complete understanding of the determinants of large-scale cooperation building in humans, though it is a key factor for the success of our species. Here I focused on the causal effect that institutional quality has on the willingness to cooperate with out-groups by studying the consequences of a major land tenure reform for cooperation with non-local anonymous strangers. I evaluate the hypothesis that awarding formalized land rights enforceable in state courts, which is a pillar of the reform that jeopardizes the village community's discretion in allocating socially-determined informal rights, increases cooperation with strangers from other villages, as measured in an incentivized multilevel public goods game.

The results support the hypothesis. Participants from villages randomly selected to have the reform implemented contribute significantly more to the account shared with anonymous peers from other villages compared to control participants. Analyses of possible mechanisms suggest that other factors that have been hypothesized as determinant of in-group favoritism, such as land-related conflicts, education investments, changes in generalized altruism, market integration or access to credit, are unlikely to have mediated the observed effects. Notably, a comparison of participants' contribution to local and non-local public goods show that experiencing the reform enlarged the scope of cooperation to non-locals without significantly affecting cooperation levels with in-groups. This evidence suggests that stronger and more impartial institutions that relax the dependence on in-group members play a pivotal role in drawing out new patterns of out-group interaction and expanding the scope of cooperation (Feinberg, Willer, & Schultz, 2014; Fu et al., 2012; Hruschka et al., 2014; Yamagishi et al., 2008). Moreover, while my research is not designed to reach conclusions regarding the long term effects of formalizing land tenure, it is possible that the observed increase in out-group cooperation will kick-start a coevolutionary process of generalized morality and well-functioning institutions, as predicted by theories of inter-generational transmission (Tabellini, 2008): formal institutions that enforce land rights more impartially will lead the population to internalize and transmit values of generalized morality, which in turn will result in increasing demand for impartial institutions. Whether my findings are capturing the initial stage of a long-term coevolutionary process, it remains a question for future research.

The article is closely related to the works of Buchan et al., 2009 who investigate the relationship between large-scale cooperation and the exposure to, and social identification with, a globalized environment. In a cross-country experimental study, the Authors document an association between living in a more globalized country or engaging in more

frequent interactions outside one's own society and an increase in the propensity to cooperate with non-local others. They show that these results are driven by social identification with the world community, a psychological mechanism that induces individuals to transcend parochial motives in cooperation decisions. My article is similar to the works of Buchan et al. in the research objective– investigating the determinants of large-scale cooperation – and in the experimental design used to elicit cooperative decisions in local and non-local domains. The findings presented here complement their approach by coupling the laboratory game with the RCT implementation of an institutional reform that dispels residual identification concerns characterizing cross-country comparisons (Hruschka et al., 2014; Pisor & Gurven, 2016). In a real-world setting it is rare to observe the random assignments of institutions and this particular feature of my research design increases the confidence regarding the direction of the causal chain. However, an advantage of the approach of Buchan et al. concerns the possibility of generalizing the research results, since their experiments involve stratified samples of participants representing the general population of six societies that are characterized by heterogeneous aggregate socio-economic conditions. Instead, the outcomes reported in my study are limited to effects observed in rural areas of a single low-income country. The extent to which the results presented here also apply to different social contexts remains a subject for further investigation.

In relation to this, it is important to emphasize that the observed effects of formalization do not apply to a random sample of Beninese villages, since the treated and control villages included in the RCT pool have volunteered to have the property reform implemented. This is a key point, since what the article tells us is that the introduction of more impartial institutions increases cooperation with out-groups, conditional on the population showing a demand for the institutional change. On the one hand, this conditionality means that the effects observed here might be different in situations of top-down institutional changes that have been supplied and enforced in the absence of local demand. What the effect of formalization in the latter situations would be is a question closely related to the larger debate concerning costs and benefits of universal (or mandatory) versus selective (or voluntary) land titling (Arrunada, 2012; Deininger & Feder, 2009) and should be established by future research. Similarly, it might be the case that the exogenously-introduced land tenure reform has also sped up the implementation of other reform interventions (of either formal or informal institutions) for which there was demand in these communities and that were not recorded in my data. Should that have happened, the results presented here would pick up the joint effects of these changes in the institutional framework. The possibility again suggests caution in taking at face value the results from a single case-study and calls for the collection of additional evidence from other settings. On the other hand, the conditionality of the results provides supporting evidence for theories hypothesizing co-evolutionary processes of institutions and large-scale cooperation (Enke, 2019; J. Henrich & Muthukrishna, 2021; Schulz et al., 2019; Tabellini, 2008). Indeed, while my research design shows that the supply of formal institutions contributes to scaling-up the cooperation networks of those having positive attitudes toward these new rules and a willingness to embrace them, a related question would be what are the conditions and reasons behind the original request for reform. In this respect, the results have points in common with the literature in anthropology and economic history that documents how a puzzling small number of human communities have been capable of enlarging the population size once a threshold of few hundred members has been reached (Casari & Tagliapietra, 2018; Forge, 1972; J. Henrich, 2020; Tuzin, 2013). My findings point toward a possible explanation for the phenomenon: once a community hits certain stages of societal development, the introduction of new formal institutions is a precondition for removing bottlenecks to the coevolution of norms of impersonal prosociality that are necessary to sustain cooperation in large groups.

While my study was not designed to evaluate the comparative performance of competing theories for in-group favoritism, it is worth discussing what can be learned from the evidence presented. The RCT design, and the fact that the data collection was performed in geographically-defined areas within a single country, make it possible to rule out that the observed increase in out-group cooperation can be ascribed to different environmental conditions (de Vliert and Evert, 2020) or exposure to pathogens (Fincher & Thornhill, 2012) between treatment groups. Similarly, the analysis does control for conflicts experienced and excludes those from being the driving force behind the effects observed in my sample (Handley & Mathew, 2020). The reform reshaped property rights in the direction of market-like institutions. However, the main results hold when controlling for several proxies for villagers' access to credit and market integration. This finding suggests that the increase in non-local cooperation has not been determined by individuals engaging more in market activities. The evidence resonates with the argument that, following behavioral channels not limited to active market participation, introducing market-like allocation rules in societies where transactions were characterized by norms of personal exchanges can favor the development of impersonal prosociality (Bowles, 1998; Hirschman, 1970). Finally, I must point out that the competing explanations for in-group favoritism which the research design here rules out, can be important determinants of cooperation in other contexts, time-spans, or ecological conditions. My future work will combine the randomized assignment of the institutional environment that has been created in Benin with sources of exogenous variations considered by other theories, in order to verify the comparative apportioning of institutional quality and other factors to shaping the extent of out-group cooperation.

In conclusion, this article suggests that supplying a society that displays demand for reforming its system of informal norms with more impartial property institutions can accelerate the scaling-up of cooperation. For instance, imagine a society that experiences an unexpected event which radically modifies existing socio-ecological conditions and calls for new forms of governance, such as environmental changes rendering sedentary agriculture feasible or technological advances making it possible to overcome the Malthusian trap (Richerson & Henrich, 2012). The creation of new institutional layers capable of capitalizing on the external shocks and enhancing societal development is neither automatic nor to be taken for granted, as historical accounts of the evolution of thousands of polities show both outstanding successes and unexpected social collapses linked to the (in-)ability to reform existing systems (Shin et al., 2020). In an era marked by global challenges such as anthropogenic climate change, threats from pandemics, and raising world economic inequalities, the results reported here suggest that institutions will play a key role in building an integrated international community capable of tackling those issues.

Data and materials availability

All data supporting the findings in this manuscript were submitted with the article as supplementary material. All custom code for data cleaning and analysis supporting the findings in this manuscript were submitted with the article as supplementary material.

Preregistration

The experiment was registered in the Registry for Randomized Controlled Trials operated by the American Economic Association before the data collection took place: see RCT ID AEARCTR-0005316.

Author contributions

M.F. is the only author and designed research, performed research, analyzed data, and wrote the paper.

Data availability

All data supporting the findings in this manuscript were submitted with the article as supple- mentary material. All custom code for data cleaning and analysis supporting the findings in this manuscript were submitted with the article as supplementary material.

Funding

This work was supported by the Marie Curie Research Grants Scheme, grant H2020-MSCA-IF-2017-789596.

Declaration of Competing Interest

The author declare that he has no competing interests.

Acknowledgements

I am grateful to the editors, Debra Lieberman and Michael Muthukrishna, and to three anonymous referees for comments and suggestions that greatly contributed to improve the paper. I also thank Benito Arrunada and Maria Bigoni for fruitful discussions, and Deo-Gracias Houndolo for support in setting up the fieldwork. The usual disclaimer applies.

Appendix A. Customary property rights and the *Plan Foncier Rural* in Benin

In Benin, as in the vast majority of the African continent, customary tenure characterized by collective property predominates in rural areas. Attempting to improve access to land, tenure security and the development of a land market, in 1993 the Beninese government developed an approach for systematic identification and registration of customary rights to parcels of agricul- tural land, the "Plan Foncier Rural" (PFR).

PFR consists of socio-land surveys at the village level to identify rights holders, their rights, and demarcate parcels boundaries. The process allows for public contestation of the pro- posed registration of rights and requires that rights holders and neighbours publicly sign survey records. Following a process of land demarcation, PFR maps that identify parcels' boundaries are produced and stored in public archives that can be freely consulted. Rightholders can also request to receive their land certificates. In the original formulation of the Rural Land Act 2007- 003 that was regulating the issuance of land rights certificates, the local administration was in charge to produce the "Certificat Foncier Rural" and automatically deliver these certificates to identified rightholders. The Certificat Foncier Rural would have then required a further admin- istrative step in order to be registered in the national cadastre and allow rightholders to obtain a land ownership title ("Titre Foncier"). The new Rural Land Law 2013-01 maintains the le- gal value of PFR-recorded rights and creates a unique ownership document, the "Certificat de Propriete Foncier," reunifies land certificates and ownership titles. We include in the regression analysis a dummy indicating whether a certificate has been delivered.

While certificates of registration do not directly confer legal ownership, nonetheless reg- istered parcels acquire a new legal status and can be transformed into land titles following a procedure that is faster and cheaper compared to standard registration of untitled land. More-over, certified rights award presumption of ownership recognized by courts and make possible selling or using registered parcels as a collateral. Thus, PFR represents a major change in the institution of property rights over land, in particular considering land is the only asset for most rural villagers.

Due to a lack of resources, de facto PFR was not implemented until 2006, when the Mil- lennium Change Account subsidized a five-year PFR program that was implemented with the help of the World Bank. This is the first, and so far unique, case of large-scale institutional re- form implemented as RCT. The reform implementation, summarized in the left panel of Fig. 1 in the main text, proceeded as follows. First, 600 eligible villages willing to implement the reform were identified. Second, a subsample of 296 villages was selected via public lottery, and PFR was actually implemented ("treatment"). Non-selected villages ("control") did not receive any intervention and, as of today, continue to have customary land rights. The random sorting of eligible villages in the two groups guarantees confounding factors are cancelled out. In the regression analysis, I control for participant's village of birth and post-treatment migration patterns.

Appendix B. Details of experimental and fieldwork proce- dures

B.1. The experiment

To elicit ingroup and outgroup cooperation, I conducted a version of a multi-level sequential public goods contribution experiment that closely follows the one used by (Buchan et al., 2009). In this game, subjects have to take two decisions. Each decision will determine earnings that are in- dependent from those generated by the other decision. At the end of the experiment, one of the two decisions is then randomly selected to determine payments. Earnings from the other decision are discarded. Subjects do not receive feedback after the first decision has been taken. For each decision, subjects are given 10 coins, and each coin is worth XOF 100 (approximately $0,17).

The first decision is similar to a standard linear public goods game. Subjects have to allocate the 10 coins between an "Individual" account and a "Village" account. Coins allocated to the Individual account are kept by the participant. Coins allocated to the Village account will be added to the coins allocated to the Village account of two other participants belonging to the same village. The total amount of coins in the Village account is then increased by the experimenter by 50% and equally divided between the three participants. Final earnings for a subject in the first decision are determined by the share of coins coming from the Village account added to the coins allocated to the Individual account.

In the second decision, subjects have to allocate the 10 coins between three accounts: an "Individual" account, a "Village" account, and a "Country" account. The Individual and Village accounts work as in decision 1. Coins allocated to the Country account by each participant are added to the coins contributed to the Country account of two participants from the same village who are also matched for the Village account and with the coins allocated to the Country account by six participants belonging to different villages in Benin. In the instructions it is specified that after Decision 1 there is a re-matching procedure, so it is unlikely that the groups of three participants contributing to the Village accounts in Decision 2 are the same as those in Decision 1. This total amount is then doubled by the experimenter before being equally divided among the nine participants. So the earnings of each participant in the second decisions are the sum of the coins allocated to the Individual account, the share of the coins coming from the Village account, and the share of the coins coming from the Country account.

This experimental design resembles that of a multilevel public goods experiment, except that subjects do not make decisions that directly affect only those in their concurrent groups in the session. More specifically, subjects decide whether to allocate coins to a Village or Country envelope, and their choice affects their own payoff. Their choice also affects the payoffs of others in future sessions. Therefore, the experiment implemented a

dynamic matching proce- dure: the past decisions from other participants in different villages were used to determine the payoffs of subjects currently taking part in the research. Specifically, the decisions of subjects were recorded during the experimental session. These decisions were then used to match pple's decisions in subsequent sessions.I used the decisions taken in pilot tests run before the experiment to determine payoffs in the initial session. The decisions of people who partici- pated in the last session of the research will be used in future experiments as starting data. This design choice was made necessary because of logistical constraints that made it impossible to run the experiments simultaneously in multiple villages. The dynamic nature of the matching procedure ensured as a consequence that the participants benefiting from one's contributions to the collective account (the "beneficiaries") and the participants whose contributions an individual benefited from (the "benefactors") were not necessarily the same. The instructions tried to convey this unusual feature of the experiment in as simple a way as possible. In particular, the instructions explicitly made it clear that (i) other people's decisions (coupled with their own decisions) would determine one's payment, and (ii) the participant's own choices would deter- mine the payments to others, depending on the group into which the subject would be mixed. An English translation of the instructions is included in the Appendix D.

B.2. Fieldwork procedures

As it is reported in the left panel of Fig. 1 in the main text, the fieldwork took place in the provinces of Mono and Couffou (in the South of the country) and Alibori and Borgou (in the North). With the support of local research assistants, I carried on the fieldwork activity col- lecting lab-in-the-field experimental and survey data on 32 villages (16 treated and 16 control) randomly selected among those included in the lottery pool.

The procedure was as follows. The day before the experiment, a research assistant met the village chief to inform him or her that the following day a team of researchers would come to the village. The village chief was requested to inform the population and gather at the convened time as many volunteers as possible for participating in the study. Participants must be resident in the village, of age, and only one member per household was allowed to participate. The day of the experiment, the research team randomly selected nine male and nine female participants among the volunteers who showed up and met these criteria. Since the fieldwork was completed during the dry season when most of the agricultural work is suspended, there were always more than 18 volunteers willing to participate. Non-selected participants were paid a show-up fee of XOF 500 (approximately $0.8) and requested to leave.

The experiment took place in a public space (usually a school or the village meeting room). The instructions were read aloud by the experimenter in a common room and the decisions were privately taken by each subject in a separate decision room. Before entering the decision room, the experimenter in private asked each participant a set of questions to verify the comprehension of the game. If the participant could not answer a question, the experimenter re-explained the instructions until the participant was able to answer correctly. The participant entered the decision room alone and physically divided the ten coins between the different envelopes, which were marked with a code. This procedure made the contribution decisions of the participants blind to the experimenter on site.

After completing the two allocation choices, participants took part to other fieldwork activities. The additional fieldwork activities involved experiments and survey questions used to test the robustness of the results and to investigate the mechanisms, as well as additional activities used for a different research project. The multilevel public goods game described in this paper was the first activity conducted during the fieldwork in each session. Participants were not informed that they would take part to the additional activities. Participants also com-pleted an incentivized risk elicitation task. Each participant was given six coins. The participant could invest part or all of the coins in a bet in which with 50% probability the amount invested will be doubled, and lost otherwise. The coins not invested were added to the participant's final payments. Moreover, participants also completed a socio-demographic survey, reporting the following information: age, gender, religion, marital status, number of family members, participation in household finance management, education, literacy, village of birth, years of residence in the village, and income. The survey asks for self-reported weekly income and, since it is well-known that in the context of developing countries income is a problematic proxy for individual wealth, a set of proxy measures: number of bedrooms in the house, whether the house has running water, concrete floor, electricity, radio or television, acres of land possessed individually, acres of land possessed by the family, whether within the household somebody owns a motorbike, a car, a bank account or a credit card.

Each fieldwork session lasted approximately three hours. Participants received as a final payment on average XOF 2800 ($4,5), roughly the equivalent to the wage earned in one and a half days of work for the median subject in our sample. We ran a session in each of the 32 village visited, for a total of 576 participants who took part to the experiment. During one session, one participant left the experiment because of feeling sick, so we actually have data relative to 575 individual choices. The data collection took place between January and March 2020.

Appendix C. Supplementary analysis

In the post-experimental survey participants were asked whether they would request a PFR certificate in case they are purchasing a parcel of land. Over 95% of the 575 respondents, both in the treated as well as in the control group, reported that they would do so. In villages where the reform was implemented, 57% of subjects report that they knew where the PFR registries are physically stored and that they can be readily consulted (The question asked to participants stated: "Do you know where PFR registries and maps are stored? Is it possible to consult them?"). Moreover, 37% of participants reported to have done so at least once in the previous nine years. These answers confirm that the registries produced by the PFR are considered reliable documents by, and that they are easily accessible for, potential users. The survey also asked participants whether they believe that, in the case of litigation concerning rights over a land parcel, having obtained PFR-registered rights would make a difference for the litigants' probability of adjudicating the case. (Participants were first described the following situation: "Two men have a conflict over property rights of a parcel of land. One man is rich; the other man is the legitimate owner and he is poor". Three questions were then asked: QI) "In front of which institution has the poorer man the highest chances of adjudicating the case?"; Possible answers: a) official state tribunal; b) customary tribunal; c) it does not matter, same probability. QII) "Imagine the same situation. However, now the poorer man has a PFR certificate which officially declares his property rights over the land parcel. Do you think the poorer man is more likely than before to keep the land if the case is litigated in front of a formal court?" Possible answers: a) yes; b) no; c) same probability, PFR certificates does not matter. QIII) "Do you think that the poorer man is more likely than before to keep the land if the case is litigated in front of a customary court?" Possible answers as in QII.) Over 90% of respondents reported to believe that having formally registered land rights increases a litigant's likelihood to adjudicate the case. Interestingly, participants reported that, in their opinion, PFR-registered rights increase the chances of success both when the case is adjudicated by a formal court and when the case is

adjudicated by a local-customary court.

The post-experimental survey collected two responses that can provide a proxy measure of trust in the government authorities. The first question asked to what extent participants trust the impartiality of the formal judiciary vis-à-vis the customary or religious dispute resolution authorities. The question states: "Two men have a dispute over the ownership of a land parcel. One man is richer and the other poorer. The poorer one is the rightful owner. However the rich man can bribe the person resolving the case with a fair amount of money, while the poorer does not have money to pay a bribe. In front of which institution is the poorer man more likely to win the case?" and the possible answers were "1) State tribunals (the formal judiciary); 2) Customary authority; 3) Religious authorities; 4) It does not make a difference". 547 of the 576 participants answered the question. 63 participants in control villages chose option 1 against 58 participants in treated. The proportion is not statistically different in the two samples (also for each of the other possible options, the proportion chosen is statistically the same in the treatment and control groups). In the experiment, the average contribution to the Country account for subjects who chose option 1 in control is 2.37 coins against 3.02 coins in treated – the difference is not statistically significant (p=0.19, t-test two-sided). The contribution to the Country account for subjects who in treated villages did not chose option 1 in control is 3.06 coins against 3.36 coins in treated – the difference is not statistically significant (p=0.25, t-test two-sided). In the second question that can be used as a proxy for trust in the government authorities, participants were requested to report their appreciation for the level of support provided by the government in a Likert scale (from 1-7, with one representing the lowest level of satisfaction). Specifically, the question stated: "Do you think the central state is helping the villagers enough relative to what it asks them to contribute?". 566 participants answered the question, with an average of 3.17 in control and 3.22 in treated – the difference is not statistically significant (p-value 0.65, t-test two-sided). If we compare the average contribution to the Country account in the experi- ment for participants who reported an evaluation smaller or equal than the median of 3, those in control contributed 2.88 coins against 3.33 coins for those in treated – the difference is not statistically significant (p=0.12, t-test two-sided). Similarly, if we compared the contribution to the Country account of those who evaluated the government action more positively, the difference across average contribution in control (2.92) vs. treated (3.46) is not statistically significant (p=0.17, ttest two-sided).

Table A1
Out-group cooperation – alternative specifications of the double-hurdle regression.

	Model 1	Model 2	Model 3	Model 4
treated	1.480**	1.464*	1.392*	1.411*
	(0.557)	(0.572)	(0.690)	(0.702)
coop-ingroup	0.423***	0.439***		
	(0.042)	(0.042)		
education		0.193		0.088
		(0.133)		(0.155)
literate		0.316		0.212
		(0.401)		(0.468)
Constant	−0.416	0.288	−1.476*	−1.951
	(1.922)	(1.908)	(1.604)	(1.567)
N.obs.	575	575	575	575

Notes: Dependent variable: Contribution to the Country account in decision 2. Double-hurdle regres- sion with standard errors robust for clustering at the village level and bootstrapping with 50 repetitions. All models include as controls: age, gender, religion, marital status, number of family members, partici- pation in household finance management, village of birth, years of residence in the village, self-reported weekly income, two dummies for villages where the PFR was extended after 2011, a dummy for par- ticipants who received a land-rights certificate, a dummy for villages in the South. Models 2 and 4 additionally controls for literacy and education levels. In Models 1 and 2 the variable *coop-ingroup* – reporting the coins contributed in a separate standard linear public goods game – and is used as control in estimating the inner hurdle; while in models 3 and 4 it is excluded by both the inner and outer hurdle estimation. In all models the explanatory variables for the outer hurdle are *age and gender*. Symbols
* * *, **, and * indicate significance at the 0,1%, 1% and 5% level, respectively.

Table A2
Out-group cooperation – adding individual and village-level controls.

	Model 1	Model 2	Model 3	Model 4
treated	1.014*	1.261*	1.001*	1.323*
	(0.478)	(0.703)	(0.487)	(0.715)
coop-ingroup	0.309***		0.307***	
	(0.087)		(0.087)	
		(0.401)		
Wealth-Ctrl	Y	Y	N	N
Village-Ctrl	N	N	Y	Y
Constant	−0.639	−0.231	−0.494*	0.189
	(1.288)	(1.945)	(1.227)	(1.840)
N.obs.	575	575	575	575

Notes: Dependent variable: Contribution to the Country account in decision 2. Double-hurdle regres- sion with standard errors robust for clustering at the village level and bootstrapping with 50 repetitions. All models include as controls: age, gender, religion, marital status, number of family members, partici- pation in household finance management, village of birth, years of residence in the village, self-reported weekly income, two dummies for villages where the PFR was extended after 2011, a dummy for par- ticipants who received a land-rights certificate, a dummy for villages in the South. Models 2 and 4 additionally controls for literacy and education levels. In Models 1 and 2 the variable *coop-ingroup* – reporting the coins contributed in a separate standard linear public goods game – and is used as control in estimating the inner hurdle; while in models 3 and 4 it is excluded by both the inner and outer hurdle estimation. In all models the explanatory variables for the outer hurdle are *age and gender*. Village Controls include: population, whether the village has a hospital or a clinic,

whether the village has a public school. Wealth Controls include: number of bedrooms, whether there is running water in the household, acres of land possessed individually, acres of land possess by the family, whether the house has concrete floor, electricity, radio or television, whether within the household somebody owns a mo- torbike, a car, a bank account or a credit card. Symbols ***, **, and * indicate significance at the 0,1%, 1% and 5% level, respectively.

Table A3
Utilitarian participants - lasso post-double-selection methodology for selection of controls.

	Model 1	Model 2
treated	1.099*	1.149*
	(0.535)	(0.545)
coop-ingroup	0.318**	0.320**
	(0.091)	(0.102)
treated-2019	−1.010**	
	(0.447)	
control-2019	0.777*	
	(0.349)	
Constant	1.128*	1.143
	(0.655)	(0.727)
N.obs.	575	503

Notes: Dependent variable: Utilitarian. Regularized post-double-selection lasso regression. Standard errors ro- bust for clustering at the village level. High-dim controls included: age, gender, religion, marital status, number of family members, participation to household finance management, literacy, years of education, whether the village of participation is also the village of birth, years of residence in the village, self-reported weekly income, incentivized measure of risk preferences population, whether the village has a hospital or a clinic, whether the village has a public school, percentage of consumed calories purchased in the market and distance from the closest market, region of the fieldwork, number of bedrooms, whether there is running water in the household, acres of land possessed individually, whether received a pfr certificate, acres of land possessed by the family, whether the house has concrete floor, electricity, radio or television, whether within the household somebody owns a motorbike, a car, a bank account or a credit card. Models 2 excludes four villages that had extended the original PFR plan. Symbols ***, **, and * indicate significance at the 0,1%, 1%, and 5% level, respectively.

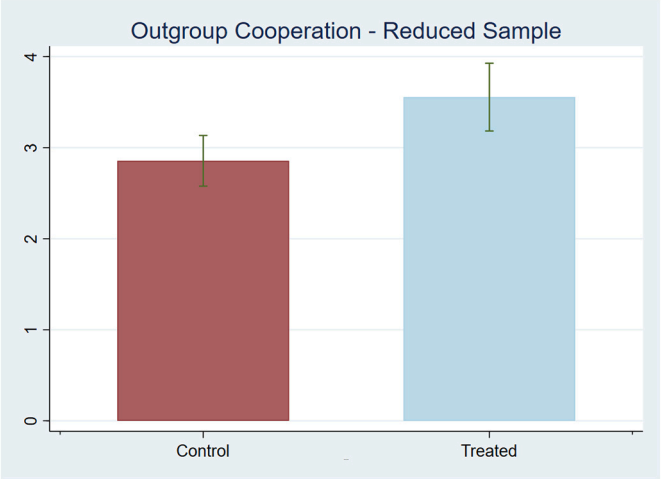

Fig. A1. Coins contributed to the country account in Decision 2, reduced sample. Vertical bars display 95% confidence intervals.

Table A4
Out-group cooperation – reduced sample.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
treated	1.043*	1.037*	1.125*	1.125*	1.043*	1.385*
	(0.486)	(0.489)	(0.551)	(0.555)	(0.486)	(0.706)
coop-ingroup	0.315***	0.313***	0.342***	0.340***		
	(0.098)	(0.096)	(0.108)	(0.106)		
education		−0.170		−0.192		−0.128
		(0.149)		(0.174)		(0.175)
literacy		0.964		0.987		0.886
		(0.538)		(0.584)		(0.675)
Constant	−0.501	−0.995	−1.434	−1.871	−0.501	−0.394
	(1.357)	(1.496)	(1.645)	(1.780)	(1.357)	(1.849)
var(e.coop2-foreign)			9.421***	9.313***		
			(1.264)	(1.251)		
N.obs.	503	503	503	503	503	503

Notes: Dependent variable: Contribution to the Country account in decision 2. Models 1 and 2: OLS regression. Models 3 and 4: Left Censored Tobit regression. Models 5 and 6: Double-hurdle regression. Standard errors robust for clustering at the village level. Models 5 and 6 implement bootstrapping with 50 repetitions. All models include as controls: age, gender, religion, marital status, number of family members, participation in household finance management, village of birth, years of residence in the village, self-reported weekly income, a dummy for participants who received a land-rights certificate, a dummy for villages in the South. Compared to Models 1, 3, and 5, Models 2,4, and 6 additionally controls for literacy and education levels. In Models 1-4 the variable *coop-ingroup* reports the coins contributed in a separate standard linear public goods game and is used as regressor. In Models 5 and 6 the explanatory variables for the outer hurdle are *coop-ingroup, age, age-squared, and gender*. Symbols ***, **, and * indicate significance at the 0,1%, 1% and 5% level, respectively.

Table A5
Out-group cooperation – controlling for conflicts experienced.

	Model 1 OLS	Model 2 OLS	Model 3 Tobit	Model 4 Tobit	Model 5 Hurdle	Model 6 Hurdle
treated	0.987*	0.972	1.063*	1.056*	1.308*	1.351*
	(0.479)	(0.483)	(0.541)	(0.547)	(0.695)	(0.708)
coop-ingroup	0.310***	0.310***	0.346***	0.346***		
	(0.087)	(0.087)	(0.096)	(0.095)	(	(
education		−0.099		−0.111		−0.141
		(0.144)		(0.171)		(0.172)
literate		0.631		0.602		0.875
		(0.498)		(0.549)		(0.661)
Constant	−0.438	−0.864	−1.508	−1.842	0.132	−0.399
	(1.234)	(1.365)	(1.489)	(1.604)	(1.763)	(1.797)
var(e.coop2-foreign)			9.350***	9.305***		
			(1.188)	(1.188)		
N.obs.	575	575	575	575	575	575

Notes: Dependent variable: Contribution to the Country account in decision 2. Models 1 and 2: OLS regression. Models 3 and 4: Left Censored Tobit regression. Models 5 and 6: Double-hurdle regression. Standard errors robust for clustering at the village level. Models 5 and 6 implement bootstrapping with 50 repetitions. All models include as controls: whether a land-related conflict was experienced in the previous nine years, age, gender, religion, marital status, number of family members, participation in household finance management, village of birth, years of residence in the village, self-reported weekly income, two dummies for villages where the PFR was extended after 2011, a dummy for participants who received a land-rights certificate, a dummy for villages in the South. Compared to Models 1, 3, and 5, Models 2,4, and 6 additionally controls for literacy and education levels. In Models 1-4 the variable *coop-ingroup* reports the coins contributed in a separate standard linear public goods game and is used as regressor. In Models 5 and 6 the explanatory variables for the outer hurdle are *coop-ingroup, age, age-squared, and gender*. Symbols ***, **, and * indicate significance at the 0,1%, 1% and 5% level, respectively.

Table A6
Out-group cooperation – controlling for market integration.

	Model 1 OLS	Model 2 OLS	Model 3 Tobit	Model 4 Tobit	Model 5 Hurdle	Model 6 Hurdle
treated	1.071*	1.044*	1.173*	1.149*	1.283*	1.342*
	(0.488)	(0.486)	(0.548)	(0.548)	(0.691)	(0.708)
coop-ingroup	0.314***	0.315***	0.351***	0.352***		
	(0.088)	(0.087)	(0.097)	(0.096)		
education		−0.064		−0.069		−0.132
		(0.148)		(0.174)		(0.186)
literacy		0.734		0.715		0.974
		(0.511)		(0.569)		(0.698)
Constant	−0.438	−0.864	−1.508	−1.842	0.359	−0.263
	(1.234)	(1.365)	(1.489)	(1.604)	(1.783)	(1.833)
var(e.coop2-foreign)			9.350***	9.305***		
			(1.188)	(1.188)		
N.obs.	575	575	575	575	575	575

Notes: Dependent variable: Contribution to the Country account in decision 2. Models 1 and 2: OLS regression. Models 3 and 4: Left Censored Tobit regression. Models 5 and 6: Double-hurdle regression. Standard errors robust for clustering at the village level. Models 5 and 6 implement bootstrapping with 50 repetitions. All models include as controls: distant from the closest market, share of total calories consumption that was purchased in the markets, age, gender, religion, marital status, number of family members, participation in household finance management, village of birth, years of residence in the village, self-reported weekly income, two dummies for villages where the PFR was extended after 2011, a dummy for participants who received a land-rights certificate, a dummy for villages in the South. Compared to Models 1, 3, and 5, Models 2,4, and 6 additionally controls for literacy and education levels. In Models 1-4 the variable *coop-ingroup* reports the coins contributed in a separate standard linear public goods game and is used as regressor. In Models 5 and 6 the explanatory variables for the outer hurdle are *coop-ingroup, age, age-squared, and gender*. Symbols ***, **, and * indicate significance at the 0,1%, 1% and 5% level, respectively.

Table A7
The *Plan Foncier Rural* and access to credit.

	Model 1	Model 2	Model 3	Model 4
treated	−0.081	0.031	−0.233	0.267
	(0.138)	(0.176)	(0.149)	(0.195)
Controls	Y	Y	Y	Y
Constant	−0.616	−0.743	−1.795***	−0.623
	(0.588)	(0.613)	(0.632)	(0.808)
N.obs.	575	575	424	424

Notes: Probit regressions. Dependent variables: Model 1 having received credit (dummy); Model 2 having the possibility to buy goods on credit (dummy);Model 3 for those who made a credit request, having received credit from a bank or financial institution; Model 4 for those who made a credit request, having received credit from somebody in the village. Standard errors robust for clustering at the village level. Controls include: age, gender, religion, marital status, number of family members, participation to household finance management, education, literacy, village of birth, years of residence in the village, self-reported weekly income, number of bedrooms, whether there is running water in the household, a dummy for villages where the PFR was extended after 2011, a dummy for villages in the South, population, whether the village has a hospital or a clinic, whether the village has a public school, percentage of consumed calories purchased in the market and distance from the closest market, acres of land possessed individually, acres of land possess by the family, whether the house has concrete floor, electricity, radio or television, whether within the household somebody owns a motorbike, a car, a bank account or a credit card. Symbols ***, **, and * indicate significance at the 0,1%, 1%, and 5% level, respectively.

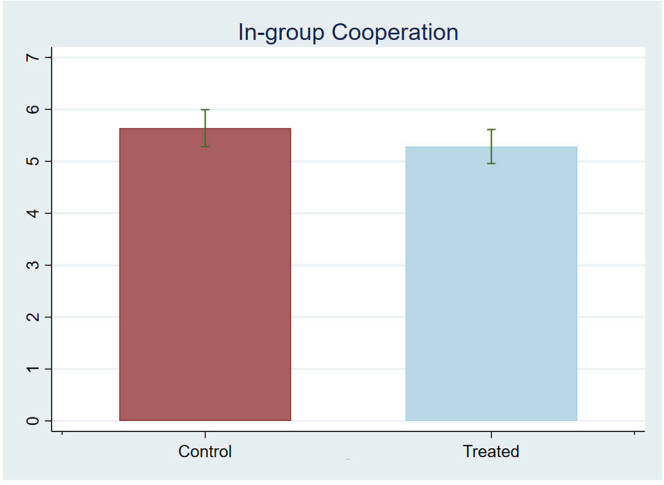

Fig. A2. Coins contributed to the Village account in Decision 1 (standard Public Goods game). Vertical bars display SE.

Table A8
In-group Cooperation-1 – contribution decisions in the standard linear public goods game.

	Model 1 OLS	Model 2 OLS	Model 3 Tobit	Model 4 Tobit	Model 5 Hurdle	Model 6 Hurdle
treated	−0.193	−0.186	−0.180	−0.171	−0.172	−0.157
	(0.469)	(0.471)	(0.475)	(0.478)	(0.544)	(0.557)
Constant	3.508**	3.583**	3.377**	3.490**	4.413**	4.542**
	(1.236)	(1.288)	(1.292)	(1.327)	(1.404)	(1.493)
var(e.coop1-local)			8.533***	8.532***		
			(0.722)	(0.722)		
N.obs.	575	575	575	575	575	575

Notes: Dependent variable: Contribution to the public good in decision 1 (the standard linear public goods game). Models 1 and 2: OLS regression. Models 3 and 4: Left Censored Tobit regression. Models 5 and 6: Double-hurdle regression. Standard errors robust for clustering at the village level. Models 5 and 6 implement bootstrapping with 50 repetitions. All models include as controls: age, gender, religion, marital status, number of family members, participation in household finance management, village of birth, years of residence in the village, self-reported weekly income, a dummy for participants who received a land-rights certificate, a dummy

for villages in the South. Compared to Models 1, 3, and 5, Models 2,4, and 6 additionally control for literacy and education levels. In Models 1-4 the variable *coop-ingroup* reports the coins contributed in a separate standard linear public goods game and is used as regressor. In Models 5 and 6 the explanatory variables for the outer hurdle are *coop-ingroup, age, age-squared, and gender*. Symbols ***, **, and * indicate significance at the 0,1%, 1%, and 5% level, respectively.

Table A9
In-group Cooperation-2 – contribution decisions to the village public good in the multilevel public goods game.

	Model 1 OLS	Model 2 OLS	Model 3 Tobit	Model 4 Tobit	Model 5 Hurdle	Model 6 Hurdle
treated	−0.236	−0.214	−0.386	−0.350	−0.364	−0.335
	(0.339)	(0.338)	(0.439)	(0.435)	(0.448)	(0.427)
coop1-local	0.220**	0.220**	0.237**	0.236**		
	(0.062)	(0.062)	(0.080)	(0.080)		
Constant	2.616***	2.869***	2.639**	3.069**	3.971***	4.305***
	(0.887)	(0.943)	(1.106)	(1.228)	(1.129)	(1.251)
var(e.coop2-local)			7.327***	7.317***		
			(0.929)	(0.920)		
N.obs.	575	575	575	575	575	575

Notes: Dependent variable: Contribution to the Village account in decision 2 (the multilevel public goods game). Models 1 and 2: OLS regression. Models 3 and 4: Left Censored Tobit regression. Models 5 and 6: Double-hurdle regression. Standard errors robust for clustering at the village level. Models 5 and 6 implement bootstrapping with 50 repetitions. All models include as controls: age, gender, religion, marital status, number of family members, participation in household finance management, village of birth, years of residence in the village, self-reported weekly income, a dummy for participants who received a land-rights certificate, a dummy for villages in the South. Compared to Models 1, 3, and 5, Models 2,4, and 6 additionally control for literacy and education levels. In Models 1-4 the variable *coop-ingroup* reports the coins contributed in a separate standard linear public goods game and is used as regressor. In Models 5 and 6 the explanatory variables for the outer hurdle are *age and gender*. Symbols ***, **, and * indicate significance at the 0,1%, 1%, and 5% level, respectively.

Table A10
Difference between contribution to village vs. country accounts in the multilevel public goods game.

	Model 1	Model 2	Model 3	Model 4
treated	−1.255	−1.210	−1.440	−1.388
	(0.729)	(0.729)	(0.840)	(0.834)
coop-ingroup	−0.088	−0.088	−0.127	−0.128
	(0.142)	(0.141)	(0.167)	(0.166)
education		0.038		0.062
		(0.214)		(0.233)
literate		−0.708		−0.885
		(0.785)		(0.890)
Constant	3.109	3.905*	3.375	4.320*
	(2.054)	(2.221)	(2.161)	(2.403)
var(e.dcoop)			18.571***	18.488***
			(3.012)	(2.965)
N.obs.	575	575	575	575

Notes: Dependent variable: Difference between contribution to the Village account and to the Country account in decision 2 (the multilevel public goods game). Models 1 and 2: OLS regression. Models 3 and 4: Left Censored Tobit regression. Standard errors robust for clustering at the village level. All models include as controls: age, gender, religion, marital status, number of family members, participa- tion in household finance management, village of birth, years of residence in the village, self-reported weekly income, a dummy for participants who received a land-rights certificate, a dummy for villages in the South. Compared to Models 1 and 3, Models 2 and 4additionally control for literacy and educa- tion levels. In Models 1-4 the variable *coop-ingroup* reports the coins contributed in a separate standard linear public goods game and is used as regressor. Symbols ***, **, and * indicate significance at the 0,1%, 1%, and 5% level, respectively.

Appendix D. Experimental instructions

Welcome to this research project. An international team of researchers is looking at the way in which people make decisions. If you pay close attention to the instructions then you could make a significant amount of money.

In this project you are going to be asked to make decisions with other people. Many people have already made their decisions and other groups are doing the same research this week. Your choices, and the choices by others, will be matched when you are finished. You will be paid in cash at the end of this research for the decisions that you and the people you have been matched with made. The same instructions are being given to other people in other villages. That is why we are reading this script. Everyone is hearing the same thing you are.

All of the decisions are similar, so please pay attention to these instructions. At the outset of each decision you will be given 10 colored tokens. Everyone will get the same materials that you get. It will be important to keep in mind that colored tokens are worth XOF 100 each to you. For other people their colored tokens also are worth XOF 100 each. Again, keep in mind that you are being matched with other people. What those people have decided to do and what you will decide to do affects how much you can make. When your decisions are submitted, we will be using our computer connection to receive information about others' choices in order to calculate each person's payments. At the end of the session, we will give you your final payment before you leave today.

You will have to make two decisions. The two decisions are independent from each other and the earnings of the two decisions are completely separated. At the end of the experiment, we will calculate the earnings for each decision and randomly choose one of the two decision. The earnings collected in this decision will be paid to you at the end of today's session. The earnings of the other decision will not be paid out. Your task is to decide

how you want to allocate your tokens between different envelopes.Here I will explain the first decision.
Decision 1

In first decision you will be given 10 tokens, and you can put your tokens into your Personal envelope or into your Village envelope. The number of tokens you put into any envelope is entirely up to you. What's the difference between the envelopes? Whatever you put into the Personal envelope is yours and will not be shared with anyone else. As mentioned before, for every colored token you put into that envelope is worth some money to you regardless of the other people's decisions. Now, what about the Village envelope? Any colored tokens that you and two other participants in today's session put into the Village envelopes will be increased by 50% by me. So, for each two coins in the Village envelope, I will add one coin. You and the other two people will get an equal share of that amount. Where do these two other people come from? They are from this village participating in today's research project. I do not know which people you will make decisions with because you will be mixed with lots of other people in order to make a group of three. All of the people you are mixed with in this decision are from the village.

You will take your decision in private inside the decision room. Before entering the decision room, I am going to pass to you your materials. You should have an envelope marked Personal and an envelope marked Village. You should have 10 red tokens (each of which are worth 100 XOF to you and everyone else). Now it is time for you to make your decision. You can put any combination of tokens into the 2 envelopes. Remember that the red tokens you put into your Personal envelope are yours and will not be divided among anyone else. Whatever you and the two other participants put into the Village envelopes will be increased by 50%. Each of you will get an equal share of that amount. Please make your decision and then place the envelopes on your box so we will know you are finished. DO NOT seal the envelopes. If you have any questions please raise your hand.

Decision 2

You have now finished the first decision. The second decision is slightly different, so please listen very carefully. In this decision you will have 10 blue tokens and 3 envelopes. Once again you will be paid XOF 100 for each blue (colored) token. In this decision you will be randomly mixed with different groups of people. The first group will be similar to the first decision. You will be mixed with two other people from this village. It is very likely that this will be two different people than the first time. The second group will be composed of 9 people. It will include the two local people, plus six other people from a rural village in Benin that is not this one.

As with the first decision, the blue tokens you put in your Personal envelope will be yours and not divided with anyone else. Second, the blue tokens you and the others put into the Village envelope will be increased by 50% and you will get a 1/3 share from the local group. Finally you have a Country envelope. The blue tokens that all 9 people put into those envelopes will be doubled. You will get an equal share of the doubled amount.

Your task is to put 10 tokens in the envelopes. You can put them in any combination that you please. Before entering the decision room, I will hand out these materials. You should get 10 blue tokens and 3 envelopes. You should have 10 blue tokens (each of which are worth XOF 100 to you and everyone else).

Now it is time for you to make your decision. You can put any combination of tokens into the 3 envelopes. Remember that the blue tokens you put into your Personal envelope are yours and will not be divided among any others. Whatever you and the two other participants in today's session put into the Village envelopes will be increased by half. Each of you will get an equal share of that amount. Whatever you put into the Country envelope will be doubled. You and 8 others will get an equal share of that amount. Please once entered in the decision room, make your decision, DO NOT seal the envelopes, and put the materials on top of your box.

Examples and comprehension checks for Decision 1

Example 1. You and the other two participants in your group put 10 tokens in the Individual envelopes and 0 in the Village envelopes.

- How many tokens do you and the other participants earn? 10.

Example 2. You and the other two participants in your group put 0 tokens in the Individual envelopes and 10 in the Village envelopes.

- How many tokens do you and the other participants earn? 15.

Example 3. You put 10 tokens in your Individual envelope and 0 in the Village envelope. The second participant in your group puts 0 tokens in their Individual envelope and 10 in the Village envelope. The third participant put 6 tokens in their Individual envelope and 4 in the Village envelope.

- How many tokens do you earn? 17.
- How many tokens does the second participant earn? 7.
- How many tokens does the third participant earn? 13.

Questions:

1) What happens when a token is put into the Village account?
2) How many people, including you, are in the Village group?
3) After every participant has made their choice and the experimenter has verified the con- tent of the Village account, how is this amount distributed among the members of the local group?

Examples and comprehension checks for Decision 2

Example 1. You and the other eight participants put 10 tokens in the Individual envelopes and 0 in the Village and Country envelopes.

- How many tokens do you and the other participants earn? 10.

Example 2. You and the other eight participants put 0 tokens in the Individual and Vil- lage envelopes, and 10 in the Country envelopes. - How many tokens do you and the other participants earn? 20.

Example 3. You and the other eight participants put 0 tokens in the Individual envelopes, 10 in the Village envelopes, and 0 in the Country envelopes.

- How many tokens do you and the other participants earn? 15.

Example 4. You put 1 token in your Individual envelope, 0 in the Village envelope, and 9 in the Country envelope. The other eight participants put 10 tokens in their Individual envelopes.

- How many tokens do you earn? 2.
- How many tokens do the two other participants who are both in your Village and Country groups earn? 12.
- How many tokens do the other six participants in your Country groups who are not in your Village group earn? 12.

Example 5. You put 0 tokens in your Individual envelope, 10 in the Village envelope, and 0 in the Country envelope. The other eight participants put 10 tokens in their Individual envelopes.

- How many tokens do you earn? 5.
- How many tokens do the two other participants who are both in your Village and Country groups earn? 15.
- How many tokens do the other six participants in your Country group who are not in your Village group earn? 10.

Questions:

1) What happens when a token is put into the Country account?
2) How many people, including you, are in your Country group?
3) How many people, including you, are in your Village group?
4) Where do the other participants from your Country group come from?

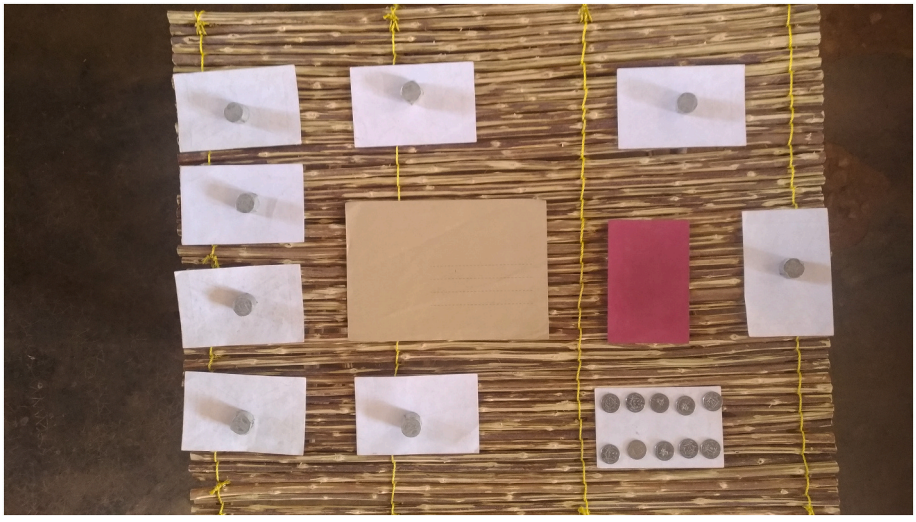

Fig. A3. Visual aid used by the experimenter during the instructions and comprehension checks. *Photo Credit: XXX.*

References

Acemoglu, D., & Robinson, J. (2019). *The narrow corridor: How nations struggle for liberty*. Penguin UK.

Alesina, A., & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature, 53*(4), 898–944.

Arrunada, B. (2012). *Institutional foundations of impersonal exchange: Theory and policy of contractual registries*. University of Chicago Press.

Balliet, D., Junhui, W., & De Dreu, C. K. W. (2014). Ingroup favoritism in coopera- tion: A meta-analysis. *Psychological Bulletin, 140*(6), 1556.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treat- ment effects after selection among high-dimensional controls. *The Review of Economic Studies, 81*(2), 608–650.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature, 442*(7105), 912–915.

Bisin, A., & Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory, 97*(2), 298–319.

Bowles, S. (1998). Endogenous preferences: The cultural consequences of markets and other economic institutions. *Journal of Economic Literature, 36*(1), 75–111.

Bowles, S., & Gintis, H. (2004). Persistent parochialism: Trust and exclusion in ethnic networks. *Journal of Economic Behavior & Organization, 55*(1), 1–23.

Bowles, S., & Gintis, H. (2013). *A cooperative species: Human reciprocity and its evolution*. Princeton University Press.

Boyd, R., & Richerson, P. (2009). Voting with your feet: Payoff biased migration and the evolution of group beneficial behavior. *Journal of Theoretical Biology, 257*(2), 331–339.

Buchan, N. R., et al. (2009). Globalization and human cooperation. *Proceedings of the National Academy of Sciences, 106*(11), 4138–4142.

Casari, M., & Tagliapietra, C. (2018). Group size in social-ecological systems. *Proceedings of the National Academy of Sciences, 115*(11), 2728–2733.

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review, 99*(1), 431–457.

Chernozhukov, V., et al. (2018). *Generic machine learning inference on heterogenous treatment effects in randomized experiments. Tech. rep. National Bureau of Economic Research*.

Deininger, K., & Feder, G. (2009). Land registration, governance, and develop- ment: Evidence and implications for policy. *The World Bank Research Observer, 24*(2), 233–266.

Devijlder, N., & Schoors, K. (2019). Land rights, local financial development and industrial activity: Evidence from Flanders (nineteenth–early twentieth century). In *Clio- metrica* (pp. 1–44).

Engel, C., & Moffatt, P. G. (2014). dhreg, xtdhreg, and bootdhreg: Commands to implement double-hurdle regression. *The Stata Journal, 14*(4), 778–797.

Enke, B. (2019). Kinship, cooperation, and the evolution of moral systems. *The Quarterly Journal of Economics, 134*(2), 953–1019.

Fabbri, M., & Dari-Mattiacci, G. (2021). The virtuous cycle of property. *Review of Economics and Statistics, 103*(3), 413–427.

Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological Science, 25*(3), 656–664.

Fiedler, S., et al. (2018). Cross-national in-group favoritism in prosocial behavior: Evidence from Latin and North America. *Judgment and Decision making, 13*(1).

Fincher, C. L., & Thornhill, R. (2012). Parasite-stress promotes in-group assortative sociality: The cases of strong family ties and heightened religiosity. *Behavioral and Brain Sciences, 35*(2), 61–79.

Forge, A. (1972). Normative factors in the settlement size of Neolithic cultivators (New Guinea). *Man, Settlement and Urbanism, 363*, 76.

Fu, F., et al. (2012). Evolution of in-group favoritism. *Scientific Reports, 2*, 460.

Galiani, S., & Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics, 94*(9–10), 700–729.

Goette, L., Huffman, D., & Meier, S. (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review, 96*(2), 212–216.

Goldstein, M., et al. (2015). *Formalizing rural land rights in West Africa: Early evidence from a randomized impact evaluation in Benin*. The World Bank.

Goldstein, M., et al. (2018). Formalization without certification? Experimental evidence on property rights and investment. *Journal of Development Economics, 132*, 57–74.
Greif, A., & Laitin, D. D. (2004). A theory of endogenous institutional change. *American Political Science Review, 98*(4), 633–652.
Habyarimana, J., et al. (2007). Why does ethnic diversity undermine public goods provision? *American Political Science Review, 101*(4), 709–725.
Handley, C., & Mathew, S. (2020). Human large-scale cooperation as a product of competition between cultural groups. *Nature Communications, 11*(1), 1–9.
Henrich, J. (2017). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.
Henrich, J. (2020). *The WEIRDest people in the world: How the west became psychologically peculiar and particularly prosperous*. New York: Farrar, Straus and Giroux.
Henrich, J., Ensminger, J., et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science, 327*(5972), 1480–1484.
Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83.
Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology, 72*, 207–240.
Hirschman, A. O. (1970). *Exit, voice, and loyalty; responses to decline in firms, organizations, and states* (p. 162). Cambridge, Mass: Harvard University Press.
Hruschka, D., Efferson, C., et al. (2014). Impartial institutions, pathogen stress and the expanding social network. *Human Nature, 25*(4), 567–579.
Hruschka, D., & Henrich, J. (2013). Institutions, parasites and the persistence of in-group preferences. *PLoS One, 8*(5).
Kaplan, H. S., et al. (2012). Risk and the evolution of human exchange. In *Proceedings of the Royal Society B: Biological Sciences* (pp. 2930–2935), 279.1740.
Lang, M., et al. (2019). Moralizing gods, impartiality and religious parochialism across 15 societies. *Proceedings of the Royal Society B, 286*(1898), 20190202.
Lowes, S., et al. (2017). The evolution of culture and institutions: Evidence from the Kuba Kingdom. *Econometrica, 85*(4), 1065–1091.
MacDonald, K. B. (2019). *Individualism and the Western liberal tradition: Evolutionary origins, history, and prospects for the future*. Kindle Direct Publishing.
Maynard Smith, J., & Szathmary, E. (1999). *The origins of life: from the birth of life to the origin of language.*.
Muthukrishna, M., Henrich, J., & Slingerland, E. (2021). Psychology as a historical science. *Annual Review of Psychology, 72*, 717–749.
Pisor, A. C., & Gurven, M. (2016). Risk buffering and resource access shape valuation of out-group strangers. *Scientific Reports, 6*(1), 1–10.
Powers, S. T., van Schaik, C. P., & Lehmann, L. (2016). How institutions shaped the last major evolutionary transition to large-scale human societies. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 371*(1687), 20150098.
Richerson, P., & Henrich, J. (2012). Tribal social instincts and the cultural evolution of institutions to solve. *Cliodynamics, 3*(1).
Romano, A., et al. (2017). Parochial trust and cooperation across 17 societies. *Proceedings of the National Academy of Sciences, 114*(48), 12702–12707.
Schulz, J., et al. (2019). The Church, intensive kinship, and global psychological variation. *Science, 366*(6466).
Seabright, P. (2010). *The company of strangers: A natural history of economic life-revised edition*. Princeton University Press.
Shin, J., et al. (2020). Scale and information-processing thresholds in Holocene social evolution. *Nature Communications, 11*(1), 1–8.
Tabellini, G. (2008). The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics, 123*(3), 905–950.
Tabellini, G. (2010). Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association, 8*(4), 677–716.
Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica: Journal of the Econometric Society*, 24–36.
Tuzin, D. (2013). *Social complexity in the making: A case study among the Arapesh of New Guinea*. Routledge.
de Vliert, V., & Evert. (2020). The global ecology of differentiation between us and them. *Nature Human Behaviour, 4*(3), 270–278.
Yamagishi, T., et al. (2008). Exchanges of group-based favours: Ingroup bias in the prisoner's dilemma game with minimal groups in Japan and New Zealand. *Asian Journal of Social Psychology, 11*(3), 196–207.
Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics, 6*(3), 299–310.