Consultative Democracy Trust

Bogliacino, Francesco and Grimalda, Gianluca and Jimenez, Laura

Universidad Nacional de Colombia, Kiel Institute for the World Economy, Universidad Jaume I, Centre for Global Cooperation Research

21 October 2017

Consultative Democracy & Trust[1]

Francesco Bogliacino[2]
Universidad Nacional de Colombia
Gianluca Grimalda
Kiel Institute for the World Economy
Universidad Jaume I
Centre for Global Cooperation Research
Laura Jiménez Lozano
Universidad Nacional de Colombia
Fondazione Franceschi

Abstract

We report experimental results from three Colombian villages concerning the impact of a voting mechanism on interpersonal trust and trustworthiness. The vote is purely consultative in that participants are asked to declare in a secret ballot the most "appropriate" plan of action for individuals involved in a "Trust Game". The plan of action that is most voted is then publicly announced. The mechanism is unbinding, as only the aggregate result of the voting is disclosed and it has no bearing on individual decisions. In spite of the strategic irrelevance of the announcement, we observe an increase in both trust and trustworthiness after the announcement is carried out, in comparison to the baseline condition where no voting takes place.

Keywords: Experiments, Trust, Voting

JEL classification: D7; C9; H4

---

[1] This paper is forthcoming on Structural Change and Economic Dynamics. This version is the pre-final version of the paper. We thank Vanessa Carrillo, Jairo Paéz and Daniel Reyes for their help during the experiments. A special thanks to Franci Beltrán, Jairo Paéz and Alfonso Peña for providing locations to run the fieldwork. Laura Jimenez thanks Fondazione "Roberto Franceschi" for the Financial Support. We thank also ASGAJUNIN for the financial support for the fieldwork in Junín. We thank the editor and an anonymous referee for their comments that helped us improving the previous version of this work. The usual disclaimer applies.
[2] Corresponding Author. Address for Correspondence: Universidad Nacional de Colombia, Carrera 30, No 45-03, Bogotá (Colombia). Email: fbogliacino@unal.edu.co

1. Introduction

A long-standing tradition in the social sciences holds that individuals are concerned exclusively with the final outcomes of a given social interaction. Nonetheless, recent approaches attribute significant relevance to the procedural aspects that accompany such outcomes (Lind and Tyler, 1988). This question is of primary importance for political scientists, who are interested in ascertaining whether democratic processes have ramifications for the legitimacy of political decisions as well as for the well-being of citizens in addition to the concrete outcomes they bring about (Olken, 2010). It is also relevant for law scholars seeking to understand how law-abiding behavior can be maximized if legal norms are perceived as procedurally fair (Tyler, 2006). Since the pioneering work by Thibaut and Walker (1975), social psychologists and, more recently, economists, have embraced the idea that individuals are concerned not only with final outcomes but also with the process leading to such outcomes. Analyses and models of procedural justice (e.g. Karni and Safra, 2002; Trautmann, 2009; Krawczyk, 2011) have started complementing those of distributive justice.

In spite of this rise of interest in processes, our knowledge of the psychological and behavioral consequences of individual preferences for procedures is still scant. In this article, we focus on a specific procedure related with consultative democracy. We do this in a controlled laboratory experiment involving an adult population from rural areas in Colombia. We ask participants to vote over the most appropriate course of action that people should hold in a given situation of interaction - a so-called Trust Game (TG) (Berg et al., 1995; Duffy et al. 2013; see section 3.2.1 for an account of the TG). The course of action that receives most votes is then publicly announced. Such an announcement is non-binding for individuals and does not change the structure of the interaction in the TG. In this respect, the vote is purely consultative. We demonstrate that holding the vote and the announcement prior to a TG significantly increases both trust and trustworthiness among individuals in comparison to not holding the vote. We conjecture that consultative voting may foster trust and trustworthiness through multiple channels. Voting can induce commitment to pro-social behavior and fosters normative expectations to comply with it. It can also stimulate

attachment to the group, thus reducing conflict. The announcement of the vote may also induce conformity. Such mechanisms are reviewed in section 2.

Existing studies have sought to measure the impact of different deliberative or political processes on individual satisfaction, through either randomized trials or surveys. Olken (2010) demonstrated with village-level randomized trials that augmenting democratic participation in social deliberation increased individuals' satisfaction with the outcome and, consequently, its legitimacy. Interestingly, the two deliberative procedures being tested do not bring about different outcomes, if not marginally. Frey and Stutzer (2005) showed evidence that Swiss citizens living in cantons where access to referenda is easier experience higher levels of satisfaction, *ceteris paribus*, than others. Anand (2001), too, used surveys to support the view that individuals are sensitive to procedural fairness. Experimental economics has demonstrated the existence of behavioral consequences of more or less fair procedures, but generally, this has been limited to their consequences on distributive outcomes (see section 2).

In this article, we focus on trust and trustworthiness because it has been theorized that inter-personal trust in strangers is at the basis of improved economic performance and institutional efficiency (Arrow, 1974; Putnam et al., 1993; Zak and Knack 2001). The relevance of trust for economic development has received extensive empirical support (Knack and Keefer, 1997; Sampson et al., 1997; La Porta et al. 1999; Zak and Knack, 2001; Guiso et al., 2006). Trust is therefore a variable of key interest for policy-makers.

No one, to the best of our knowledge, has studied the impact of a consultative voting mechanism on trust and trustworthiness. This is a significant gap in the literature for at least two reasons. On the one hand, in developing and developed countries, local consultation of communities for projects is quite a standard institutional tool, for both empowerment and establishment of social norms. It is then important to appreciate its effects in a controlled setting to better understand its impact in reality. On the other hand, a strand of the experimental literature stresses the positive advantages of introducing efficiency-enhancing institutions endogenously, that is, by vote. For instance, Tyran and Feld (2006) showed that voting for law backed by non-deterrent sanctions (so-called mild law) increases cooperation, whereas exogenously imposing the same law bears no effects. Likewise, Rauchdobler et al.

(2010) found both signaling and commitment effects from voting on minimal thresholds in public good games, and again no effects when the same law is imposed exogenously. In this literature, the newly introduced institutions alter the game payoffs. This is however not the case for consultative democracy. Our study is therefore novel in ascertaining the effect of voting when the incentives in the game are not affected by voting.

We also believe that conducting experiments in rural areas of developing countries is particularly valuable. We know from growth theory literature that weak institutions lead to multiple equilibria, which are Pareto-ranked (Bidner and Francois, 2010). Colombian institutions, especially in rural areas, are clearly weak. It is real-life decisions taken by low-educated, low-income people from rural areas – as the participants in our sample - that may have a direct impact in fostering entrepreneurship, alleviating poverty, and triggering development. We thus believe that this sample is particularly relevant to investigate the root causes of under-development of the region. As we will discuss in section 4, our rural sample shows level of trust different from urban samples from various Latin American countries (Cardenas et al., 2013). This difference is a further element of interest of our contribution.

More generally, we also agree with social scientists criticizing the excessive reliance on "WEIRD" samples - where "WEIRD" stands for Western, Educated people from Industrialized, Rich, and Democratic countries, as per the acronym used in Henrich et al. (2010). Harrison and List (2004) also suggest that the lab is a "sterile environment" to gather "general" evidence, and advocate the conduction of, among other settings, "lab-in-the-field" studies. These insights confirm the necessity of research with "non-WEIRD" samples in field experiments, as we do with the present study.

This article proceeds as follows. Section 2 reviews the literature on the TG and how different institutional mechanisms can foster pro-social behavior. Section 3 presents our experimental design. Section 4 illustrates the results. Section 5 discusses results and Section 6 concludes. The Appendix includes the full protocol (instructions, decision sheets and extract of the questionnaire) translated into English.

2 Empirical background

2.1 Individual-level factors of trust and motivations

The TG was introduced by Berg et al. (1995) to capture the key aspects of a situation involving trust. In the TG, a first-mover can send some money to a second-mover, who then decides how much money to send back. The money transferred from the first to the second mover is multiplied by a factor greater than one to capture the efficiency gains that can be reaped carrying out an economic transaction. The first mover is nonetheless exposed to the risk of losing the money she sends, if the second mover returns less than the amount sent. The first-mover is *vulnerable* to be taken advantage of by the second-mover, and this situation of vulnerability is indeed a key constitutive element of trust (Kollock, 1994; Hardin, 2002). Berg et al. (1995) focused on two scenarios, one in which the results of previous interactions were revealed to subjects, one where this was not the case. In both cases, the TG was played "one-shot" so no reputation can be built.

Several motivations may affect behavior. A first-mover may be motivated by altruism and concerns for efficiency, if she attaches value to increasing the pie to be distributed between the pair, and if the second mover's utility enters as a utility-enhancing argument into her own utility. A first-mover transfer to the second-mover may be simply motivated by self-interest, if she thinks that the second-mover will return more than what has been sent. This is particularly the case if the first-mover thinks that a norm of reciprocity will somehow urge the second-mover to share part of the profit that the first-mover made possible with her transfer. The second-mover decision is non-strategic, because she suffers no monetary consequences after her action. The second-mover decision to return a positive amount of money may again be driven by pure altruism, i.e. by the desire to share part of her payoff with somebody else. Inequality aversion may be an additional motivation, as any positive transfer will increase the second-mover's money endowment above the first mover's one. An additional component of the second-mover's behavior is reciprocity. That is the desire to return some money to the first-mover *because* of the perceived "kindness" of the action carried out by the first-mover. In fact, reciprocity is one of the first motivations that received attention in behavioral economics (Rabin, 1993) and has been the subject of extensive theoretical modelling (Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006).

Clearly, these motivations often overlap with each other and disentangling them requires suitable experimental design. By comparing experimental behavior in a TG and in versions of a Dictator Game that have the same monetary structure of incentives as the TG, but not the same strategic or reciprocity structure, Cox (2004) was able to tell apart some of the motivations spelled out above. He found that about half of first-mover behavior can be accounted for by altruism and efficiency concerns, and the remaining half by trust proper. The latter is defined as the belief that a first-mover investment in a TG will bear a positive return, for the transfer in excess of the amount that would be sent for altruistic reasons. Betrayal aversion (Bohnet et al., 2008) can also affect the preferences over the amount being sent. Likewise, about half of second-mover return rates appear to be motivated by altruism and inequality aversion, and the remaining half by reciprocity proper. The latter is defined as the desire to reward kind actions with costly transfers, in excess of what pure altruism and inequality aversion would command.

Other studies demonstrated that trust is correlated with a wide range of individual traits. Age (Belli et al. 2012), hormone levels (Ball et al. 2013), neuropeptide oxytocin (Klackl et al. 2013), gender (Buchan et al. 2008), presence of personality disorders (Franzen et al. 2012), and exposure to environmental degradation (Karapetyan and d'Adda, 2014), can all affect the amount transferred by a first mover in a TG. A hypothesis that has been investigated concerns that risk attitudes may affect trust (Fehr, 2009; Karlan, 2005). This seems plausible as the decision to transfer money to the second-mover embeds a component of risk similar to a financial investment. Surprisingly, most experiments failed to prove that risk attitudes predict first-mover's behavior in the TG (Eckel and Wilson, 2004; Houser et al. 2010). This may be due, however, to confounding effects in the way risk is measured in a "financial" context as compared to an inter-personal strategic context. Betrayal aversion may of course also prevent finding a direct relationship.

2.2 Institutional factors affecting trust

As mentioned in the introduction, trust is deemed as a fundamental component of economic prosperity. Several studies have studied how modifications of the TG rules may increase efficiency, with an eye to see which institutions may increase pro-social behavior in societies.

2.2.1 Commitment to promises and normative expectations

The possibility to make promises has been shown to increase pro-social behavior such as trust (Ellingsen and Johannesson, 2004). Two, possibly complementary mechanisms, are at work. On the one hand, a commitment effect induces individuals to stick with their promises. This may be due to an intrinsic desire to be consistent with a previously announced behavior (McClennen, 1990; Cialdini, 1993), or to the willingness to comply with a social norm prescribing that promises should be kept (Ostrom et al., 1992). Failing to commit with one's promise may stir emotions of disappointment for having failed to live up with one's own moral standards and social identity (Akerlof and Kranton, 2000), or internal reference points (Kőszegi and Rabin, 2006; Bogliacino and Ortoleva, 2015).

While commitment exclusively concerns the actions of the individual making the promise, people may also be concerned with the consequences that one's promise has for the counterpart. In particular, individuals may be concerned with disappointing people's expectations, once a promise has been made. The notion of normative expectations (Sugden, 1986; Cialdini et al., 1990; Bicchieri and Xiao, 2009) stresses that others' expectations provide individuals with a reason for action that has both motivational and moral cogency. Normative expectations may have an emotional content, because betraying others' expectations may cause a sense of guilt (Battigalli and Dufwenberg, 2007).

Vanberg (2008) proposed a design to disentangle these two effects. First-movers were either kept in the same interaction with the second-movers to whom they had made a promise, or reassigned to new second-movers who did not receive any promise from the first-mover, but whose expectations were communicated to the first-mover. The results confirm that promises increases trust, with the commitment effect being the main underlying motivation. Knowing the second-mover expectations did not increase trust in comparison with the baseline. The introduction of an intermediary increases investment if there is transparency on the exchange between the intermediary and the trustee, but not between the former and the investor (Rietz et al. 2013).

2.2.2 Announcement

The announcement of a course of action, by an institution or an individual in position of authority, may also increase pro-sociality. This may be the case simply because individuals

are inclined to imitate others (e.g. Bargh et al., 1996), or because the announcement induces convergence of beliefs over a common course of action. In fact, the announcement that a course of action is "appropriate" to a certain situation, or is "commendable", may affect individual beliefs on what others expected from one self, thus creating normative expectations (see section 2.2.1). The literature has contrasted *empirical* expectations – i.e. expectations triggered by behavior actually performed by others (Sugden, 1986; Cialdini et al., 1990; Bicchieri and Xiao, 2009) – with *moral* expectations - i.e. expectations triggered by the communication of "appropriate" or "moral" behavior (Bicchieri and Xiao, 2009; Bracht and Feltovich, 2009). Although both have behavioral relevance, empirical expectations have been proved to carry stronger motivational cogency than moral expectations. Berg et al. (1995) found that informing participants of other participants' previous behavior significantly increased second-movers' return rates, but left unaffected first-movers' transfers.

2.2.3 Communication

One of the strongest finding in social psychology is that letting people communicate with one another before undertaking some common tasks increases cooperation (Dawes, 1980; Ostrom and Walker, 1991; Balliet, 2010). In particular, cheap talk was shown to increase trust and trustworthiness (Bracht and Feltovich, 2009), contrary to Nash equilibrium predictions that unbinding communication should not change strategies by self-interested rational agents. The reason for this effect may lie with communication or other forms of interaction increasing attachment to the group (Baumeister et al., 1995; Korsgaard et al., 1995). This sense of attachment induces individuals to think of themselves as part of the group, rather than as isolated agents (Bacharach, 1999), and to replace group goals for individual goals (Kramer and Brewer, 1984). Another way to look at this phenomenon is to think that communication reduces social distance, which is known to hamper trust formation (Lei and Vesely, 2010).

2.2.3 Procedural fairness and "voice"

Another strand of literature stresses how individuals' concerns for procedural fairness may affect behavior, possibly favoring cooperation. A seminal study showed that procedural fairness - intended as equal chances to achieve unequal outcomes - was a substitute for

outcome fairness - intended as outcome equality (Bolton et al., 2005). Other studies replicated this result, although they showed that equality of opportunity was not a full substitute for equality of outcomes (Becker and Miller, 2009; Krawczyk and Le Lec, 2010). Bolton, Brandts and Ockenfeld (2005) also found receivers' opposition to procedures that were strongly biased in proposers' favor. Karni et al. (2008) showed that about 50% of their sample was willing to pay to have fairer procedures to determine final payments. Grimalda et al. (2016) showed that an even marginal increase in the fairness of a given procedure induced a large rise in pro-sociality.

Based on this evidence, one may argue that implementing procedures that are perceived as being fairer by individuals, may induce higher pro-sociality. Voting over a certain social interaction may be one of such mechanisms, because it permits the expression of one's "voice" over a situation (Anand, 2001; Frey and Stutzer, 2005), and may also strengthen sentiments of attachment to the group (see section 2.2.3).

2.2.5 Sanctions

Even if they will not play a part in our experiment, other institutional mechanisms have been proved to increase pro-social behavior. A well-established result in cooperation games is that allowing group members to inflict sanctions on others can massively increase cooperation. This is the case when punishment is administered both by other group members – as in so-called second-party punishment (Fehr and Gächter, 2002) – and by by-standers – as in so-called third-party punishment (Fehr and Fischbacher, 2004). This result, however, is culture-specific (Gächter et al., 2010), because anti-social punishment can seriously hamper the effectiveness of punishment (Herrmann et al., 2008). Within cultures where endogenous sanctions are effective in raising cooperation levels, one may expect the same result to hold for TGs. TGs have a similar incentive structure to cooperation games, because individual self-interest and social efficiency diverge in both games. Charness et al. (2008) showed that this is indeed the case, as third party punishment applied to second-movers significantly improved trustworthiness. A "crowding out" of intrinsic motivation can nonetheless reduce this effect (Fehr and Rockenbach, 2002).

2.2.6 Reputation

Another well-established result is that repeated interactions increase pro-sociality in comparison with single interactions. Interestingly, this is not only the case in an infinitely repeated interaction, where cooperation is supported by the folk theorem (Fudenberg and Maskin, 1986), but extends to finitely repeated games. In the latter context, Boero et al. (2009) and Dubois et al. (2012) detected a significant effect of reputation formation as a channel to foster trust and trustworthiness. Charness et al. (2011) tested a reputation mechanism through releasing information on past behavior and detected a significant increase of trust. Repeated interaction positively affected sending and return rates in Cochard et al. (2004), and this may be connected to dynamic strategic effects. Voluntary leadership was shown to increase contribution in Kleine et al. (2014). Voluntary disclosure of the decision can increase or not investment depending on the prior probability held by counterparts over the trustworthiness of the sender who discloses the information (Lunawat, 2013).

3. Experimental design

3.1 Ethnographic account of the sample

Participants in our research were adults recruited in the three small villages of Junín, Cota, and La Mesa, for a total of 91 subjects. Although the sample is small, it reflects actual calculation with G*power for a Wilcoxon signed rank test (which will be our main statistical test), using 5% and 90% as threshold level of significance and power, and an effect size (*dz*) of 1.5.

Recruitment was performed through announcements on local radios, in churches, at community meetings and approaching people on the street. These villages are located in a rural area of Cundinamarca - the region surrounding Colombia's capital, Bogotá. Junín has around 8000 inhabitants, the rural part of Cota around 10000, and the rural part of La Mesa around 13000. In all cases the population is very disperse and culturally homogeneous, with negligible presence of ethnic minorities (which are not sampled). All the three villages are relatively close to Bogotá and are currently in similar conditions of security, without

significant exposure to violence of the population. Most of the population is employed in agriculture or in the service sector in the center of the village.

Summary statistics for the sample demographics characteristics are reported in Table 1. Although some differences across locations exist - e.g. educational level in La Mesa is higher than in the other two villages -, in all three villages the sample of participants shows striking differences with respect to university student samples. Half of the sample has at most secondary education. People with low socio-economic status are amply represented in our sample, in contrast with university student samples. Although we cannot make claim of representativeness of our sample for the broader population, we believe that it is important to go beyond the usual convenience student sample used in most experimental research.[3]

We have a reliable measure of socio-economic status coming from the rating of a participant's dwelling (see Question 7 in the Questionnaire reported in the Appendix). This rating is performed by local village councils and offers a realistic estimation of the value of the dwelling and, as such, of the resident's standard of living. The rating is usually correlated with income (Bonilla et al., 2014). People are generally available to mention this rating, so this measure does not suffer from false reporting or lack of reporting that is commonly associated with reporting one's income level.

Table 1: Descriptive statistics

	Total	**Junín**	**Cota**	**La Mesa**	**Independence across village (χ^2; p value)**
Sex (%)					(6.99; p=0.03)
Male	53.3	55.5	34.4	67.6	
Female	46.6	44.4	65.5	32.3	
Age (years)	39.2	41.4	48	29.6	(102.35; p=0.06)
Socioeconomic status (%)					(26.20; p=0.001)
(1-2)	62.9	70.4	93.9	32.5	
(3)	27	22.2	7.1	47.0	
(4 or more)	10.1	7.4	0	20.5	
Household size (people)	4.9	6	4.2	4.6	(22.68; p=0.20)
Education (%)					(26.71; p=0.003)
Primary	21.8	28	37.9	2.9	

[3] All participants are stable residents of the villages. Qualitative evidence from the de-briefing after the sessions suggests that we reached typical villages inhabitants, well aware of the problems of the community, especially with respect to the rural area. Although not conclusive, this evidence seems encouraging in terms of external validity.

Secondary	27.5	44	20.6	20.5	
At least some university	28.6	12	17.1	49.9	
Technical schools	21.8	16	20.6	26.4	
Occupation (%)					(19.80; p=0.003)
Abstract task (ISCO 1-3, 6)	26.1	20	48	14.7	
Manual task (ISCO 4-5, 9-10)	22.6	36	20	14.7	
Routine task (ISCO 7-8)	7.1	16	-	5.8	
Retired, unemployed, student	44	28	32	64.7	
Observations	91	27	29	34	

The motivation for this investigation came from an ongoing collective project in Junín, whose aim is a better management of water supplies. These are affected by free-riding problems in raising voluntary contributions. Water for domestic use in Junín is provided through two small rivers. Negligent practices by local farmers in raising cattle pose a serious threat to the quality of water, because of the threat of contamination. Similar problems of water management characterize many small villages in the area. The population was thus involved in a collective project to raise funds from village residents. Such funds are used to pay local farmers to build enclosures that prevent the cattle from accessing the rivers. This project has been recently implemented and its constitutive elements resemble a TG. Residents transfer money to farmers through a local NGO. Farmers may either "pay back" building the enclosures, or "keep the money" for themselves without taking any action benefiting the donors. It is worth stressing that the "pay back" action by farmers is not enforceable by law. Thus, it relies exclusively on their good will. The TG played in our experiments may thus have been construed as close to a real-life situation.

3.2 The experimental decisions and procedures

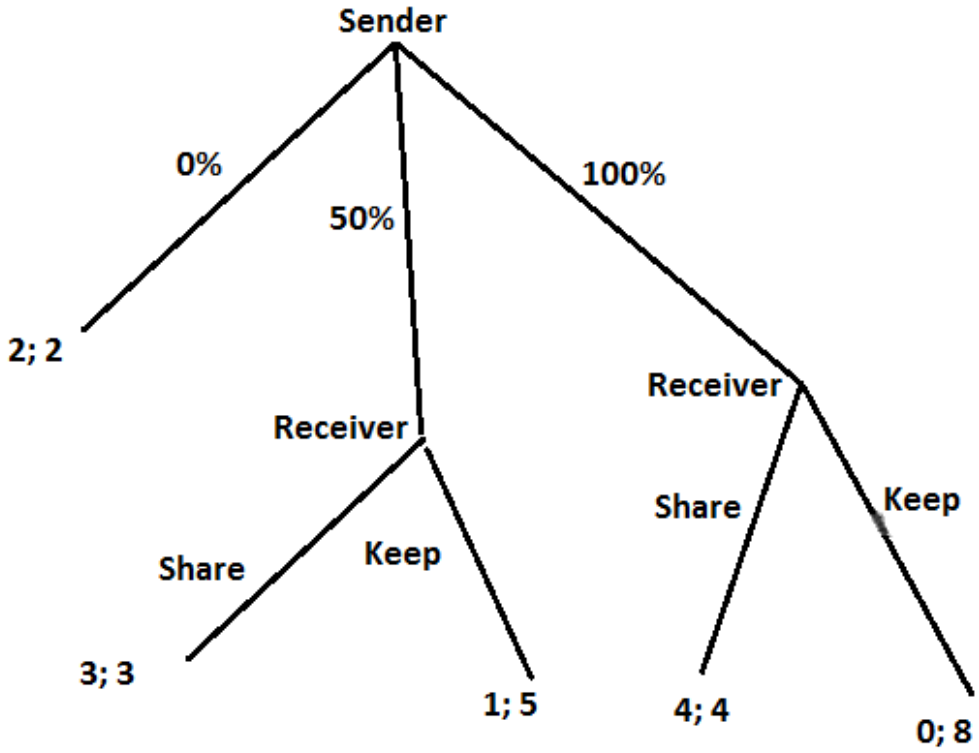
3.2.1 The stage game

Our stage game was a simplified version of Berg et al. (1995) TG, similar to Charness et al. (2011) and Ermisch et al. (2009). The sender and the receiver were given two tokens each at the beginning of each interaction; the sender then decided whether to transfer 0, 1, or 2 tokens to the receiver. We call this variable Amount Sent (AS). The amount transferred was tripled by the researcher. The receiver then decided whether to keep all the tokens in her possession, or to share them equally with the sender. We applied the strategy method, therefore the receiver had to indicate her Return Rate (RR) for both cases of the sender

sending 1 token and the sender sending 2 tokens. The extended form of the game is illustrated in Figure 1, with payoffs given in tokens. Each token was worth 4000 Colombian Pesos (COP). The hourly minimum wage is slightly above 2000 COP (around 1.3 USD), so stakes were certainly meaningful for participants. We chose this version of the game mainly for its greater simplicity in comparison to versions where senders have many more options.

It is well known that the Nash equilibrium of the game prescribes the first-mover to send nothing. A self-interested second-mover would return nothing to the first-mover, who is then better off by sending zero tokens. This course of action is however rarely played out in a TG. The social optimum of the game is instead for the first-mover to send two tokens and, assuming utility functions concave in payoffs, for the second-mover to split the pie evenly.

Figure 1: The trust game in extended form.



3.2.2 Experimental procedures and treatments

Participants received a randomly assigned ID number at the beginning of the session. We took care that half of participants received an odd number and half an even number. Once they were seated, the lead experimenter informed that subjects would take a series of decisions, and that they would be paid for only one randomly selected decision at the end of the session. Neither the number of decisions nor their nature was specified. The lead experimenter then illustrated the stage game on a blackboard. Visual representations of the interactions were also handed out (see Figure 3 in the Appendix for an example).

Comprehension was checked through four multiple-choice questions. The exactness of the answers was checked individually and the number of errors was recorded. In case of errors, subjects were asked to try to answer again and call the researchers when finished. If answers were still incorrect after this second attempt, the interaction was illustrated once

more individually and then subjects were asked to answer once more. No subject failed this second trial of questions.

At the end of the comprehension check, it was announced that everyone had answered correctly. The role of sender and receiver was then assigned. A random draw was carried out assigning either people holding an odd ID number or an even ID number the sender role. The other group was assigned the receiver role. Both senders and receivers were asked to submit their decision regarding the stage game on a decision sheet. Receivers made their decision through the strategy method. That is, they had to answer whether they wanted to keep all the tokens or half of the tokens, conditionally on the sender transferring one or two tokens. When everyone had made their choices, the decision sheets were collected.

Subjects were then told that the second decision consisted in another TG, where roles would be swapped. Those who acted as receivers (senders) in the first decision acted as senders (receivers) in the second decision. New pairs were formed. Our matching algorithm ensured that everyone was paired with a different player from the first decision. This was publicly announced. This prevented direct reciprocity to play a part in the second, as well as in any, decision. We refer to the first and second decision just described as 'Round 1' of the experiment.

After the second decision was completed, procedures differed in the Baseline and in the Voting condition. In the Baseline condition, subjects were asked to play two more TGs, one in the sender role and one in the receiver role, replicating the procedures followed for Round 1. The newly formed pairs differed from those of the previous decisions. In the Voting condition, this second set of decisions was preceded by a vote. Each subject was asked to indicate on a sheet which actions she considered the most appropriate for participants to perform[4, 5]. Subjects were asked to vote for all the three decisions that subjects would make–

[4]The text of the relative instructions was as follows: "Please, indicate how many tokens you would consider appropriate for the sender to send to the receiver" "For each scenario, please indicate if you consider appropriate that the receiver send some tokens." See the Appendix for the full version of the instructions.

[5] The choice to have a consultation on the play itself is motivated by the need to connect the problem of social norm with a direct observable behaviour. Moreover, we follow standard practice in experimental design, avoiding tasks that may generate problem of comprehension, induce some source of bias and introduce further contextual drivers of behaviour that may violate dominance axiom according to induced utility (Friedman and Sunder, 1994; Smith, 1976).

namely, the sender's transferred amount, the receiver's return when the sender transferred one token, and the receiver's return if the sender transferred two tokens. The actions that received most votes were then publicly announced. It was explained that the subsequent decisions continued to be anonymous, and that the consultation was not binding. For each village, we randomly assigned one out of three sessions to the Baseline condition and the remaining two to the Voting condition. After the Voting, the third and fourth decisions in the Voting condition followed the same procedures as the Baseline. We refer to the third and fourth decisions as 'Round 2'. The decision structure for Baseline and Voting conditions are represented in Table 2.

Table 2: The Experimental Design: Timing of Decisions

Baseline		
Round 1: TG		Round 2: TG
Voting condition		
Round 1: TG	Consultation + Public announcement of voting	Round 2: TG

At the end of the session the decision determining the payments was randomly drawn. Payments were prepared while participants completed a questionnaire. The overall experiment took up to two hours, with minimal variation. Average payment was 13275 COP (around 6.50 USD) with standard deviation 5020 COP (2.6 USD).

3.3 Discussion of our research hypothesis

Our main research question is whether introducing a consultation phase increases pro-social behavior, in particular trust and trustworthiness. We expect decisions to be similar between Baseline and Voting conditions in Round 1. Our key hypothesis is that the introduction of the consultation stage in the Voting condition increases both the amounts sent and the amounts returned in the Voting condition in comparison with the Baseline condition in Round 2.

We reviewed a variety of mechanisms as to why this should be the case in Section 2.2. The act of voting may create a commitment effect, i.e. a desire to stick with the promised behavior as expressed in the voting (section 2.2.1). Alternatively, the announcement of the vote may create a set of convergent expectations on the course of action to be followed. Such expectations may have had a normative character for individuals, as argued in section 2.2.1, because betraying others' expectations may induce guilt. This was particularly the case as the most voted norm was presented as the "most appropriate behavior" to take in the choice in Round 2, and because in all of our experimental sessions the most voted norm was the social optimum (section 3.2.1, 4.1). Announcements may also exert a positive effect of cooperation regardless of their normative content (section 2.2.2).

Furthermore, the act of voting may have an effect in strengthening attachment and identification with the group, in analogy with findings from the literature that sees communication among group members improving cooperation. Even if no verbal communication was permitted in the experiment, democratic voting may be seen as a form of group interaction that fostered attachment to the group. This makes possible the replacement of group objectives for individual objectives (section 2.2.3). Finally, the introduction of voting may increase the perception of the fairness over the overall interaction. The "voice" that voting gives to people may have a role in reducing conflict and increasing pro-sociality (section 2.2.4).

4. Results

4.1 Exogeneity of treatment

Firstly, we assess the exogeneity of the treatment. We can test the balancing of socio-demographic variables across the two conditions (respectively we have 62 participants in the Voting treatment and 29 in Baseline). We perform a two-tailed Mann-Whitney-Wilcoxon test. The null hypothesis that observations come from the same distribution is not rejected at the 10% level with regards to sex ($z = 0.21$; $p = 0.83$); age ($z = 1.46$; $p = 0.14$); socio-economic status ($z = 0.79$; $p = 0.42$); educational level ($z = 1.13$; $p = 0.25$), occupation ($z = -0.94$; $p = 0.34$). The same holds for the number of errors in the first trial of the

comprehension check, which can be taken as a measure of cognitive capabilities ($z = 1.63$; $p = 0.10$).

4.2 Results about decisions

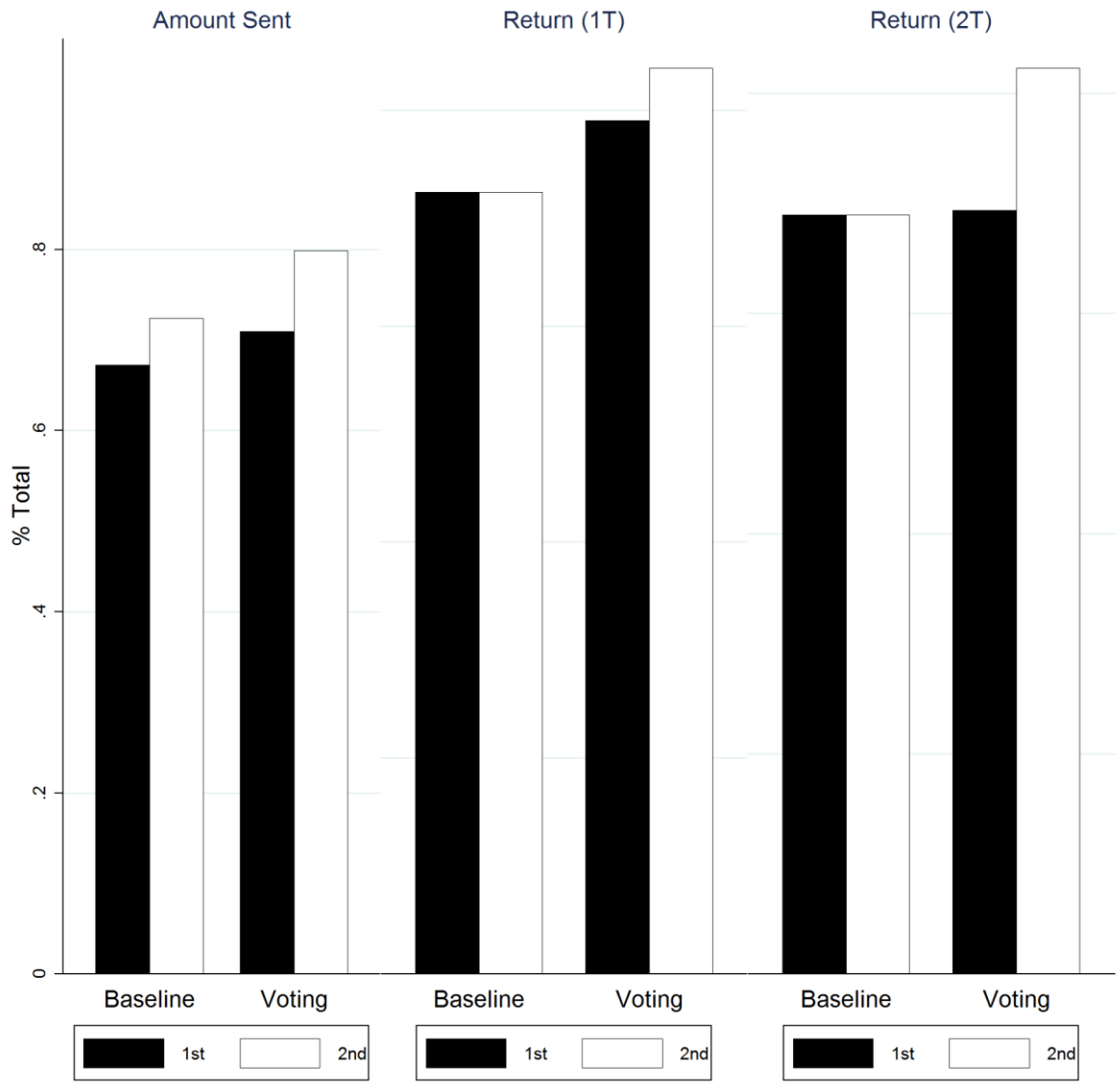
Table 3 and Figure 2 report the distribution of decisions per round and treatment condition. We do not find any significant difference between Baseline and Voting conditions in the first round. We test for the null hypothesis that observations come from the same distribution through a Wilcoxon Mann-Whitney rank sum test. The null hypothesis is not rejected for amount sent ($z = -0.79$; $p = 0.42$), for return if AS=1 ($z = -0.69$; $p = 0.48$), and for return if AS=2 ($z = -0.03$; $p = 0.97$). Moreover, we do not find any significant differences between Decisions 1 and 2 in Round 1 or Decisions 3 and 4 in Round 2. Hence, the order in which decisions were taken within each round does not appear to matter for subjects.[6]

[6] For each round we perform a Mann-Whitney-Wilcoxon rank sum test using the order with which a decision was taken as the independent variable. As far as Round 1 is concerned, the test yields $z=-0.638$ ($p=0.52$) for the Amount Sent $z=0.19$ ($p=0.84$) for Return if 1 token is sent, $z=0.38$ ($p=0.70$) for Return if 2 tokens are sent. As for Round 2, the statistics for the same variables are respectively, $z=-1.03$ ($p=0.29$), $z=-0.05$ ($p=0.95$), and $z=-1.56$ ($p=0.11$).

Table 3: Decisions per round and treatment

	Baseline		Voting	
	First Round	**Second Round**	**First Round**	**Second Round**
Amount sent:				
Average	1.34 (0.55)	1.44 (0.57)	1.41 (0.66)	1.59 (0.55)
Percentage sending 0 tokens	3.33	3.33	9.67	3.22
Percentage sending 1 token	60	50	38.7	33.87
Percentage sending 2 tokens	36.66	46.66	51.61	62.9
Return rate, if 1 token sent	0.72 (0.45)	0.72 (0.45)	0.79 (0.41)	0.83 (0.37)
Return rate, if 2 tokens sent	0.68 (0.47)	0.68 (0.47)	0.69 (0.46)	0.82 (0.38)
Observations	29	29	62	62

Figure 2: Average Sending and Return rates, by round and treatment.



Note: Amount Sent identifies the amount sent by first-movers as a share of the total tokens. Return (1T) is the share of second-movers who divided the pie evenly if one token was sent; Return (2T) is the share of second-movers who decided to divide evenly the pie if two token were sent; 1st and 2nd identify Round 1 and Round 2, respectively.

Overall, levels of both trust and trustworthiness appear large. Less than 10% of senders send nothing, and nearly half of the senders send the whole endowment. Less than 15% of

receivers do not share the surplus evenly with the sender. This is in stark contrast with, e.g. Cardenas et al. (2013), who find much lower levels of trust (the average amount sent is 34% in their experiment), and trustworthiness (the average return rate is 19.2% in their experiment). This difference may be caused by the different structure of the game (they implement a larger set of strategies for both trustor and trustee), or by differences between rural and urban environments.

As for Round 2, Table 4 reports the voting percentages in the Voting treatment. As far as 'Amount Sent' is concerned, voters expressed a marginal preference for sending 2 tokens over 1 tokens, while only a minority voted for sending 0 tokens as the most appropriate behavior. An overwhelming majority indicated that returning half of the endowment was the appropriate course of action when acting as a receiver in both of the possible cases.

Two clarifications are in order. First of all, strategic motives by participants to influence direct counterparts through votes are minimized by the fact that only results and not individual votes are revealed (this is different from direct communication) and by the fact that the chance of being the pivotal voter is negligible. Secondly, one may wonder whether there is some experimenter effect, especially in the form of Social Desirability Bias (SDB), given that this choice is not incentivized. We think that the problem is minimized by the fact that the post vote choice is monetarily incentivized. Moreover, the double blind procedure implemented in the referendum also minimizes the risk of SDB. Empirically, education and age are considered the two main triggers of SDB (Nederhof, 1985; Palhaus, 1984). However, vote is not statistically different across neither age (vote for send choice, $\chi2 = 76.20$, $p=0.18$; vote for return, 1 token, $\chi2 = 44.13$, $p=0.09$; vote for return, 2 tokens, $\chi2 = 41.25$, $p=0.15$) nor education (vote for send choice, $\chi2 = 11.45$, $p=0.32$; vote for return, 1 token, $\chi2 = 5.69$, $p=0.33$; vote for return, 2 tokens, $\chi2 = 3.44$, $p=0.63$). Finally, we believe that the SDB would be more likely to occur were the announcement an expression of the experimenter's prescription. Since in our experiment the announcement is on the contrary the indication of the subjects' own deliberation, we believe that SDB did not play a significant role in our experiment.[7]

[7] Previous acquaintance among the participants is not an issue either. First of all, we made sure, recruiting at different points in the village for every session, that the risk of previous knowledge is minimized. Secondly, the

Table 4: Voting decisions

Voting for Sender's choice	
0	9.68%
1	41.94%
2	48.39%
Vote for Return Rate, if 1 token is sent	
Return half (%)	85.48%
Return nothing (%)	14.52%
Vote for Return Rate, if 2 tokens are sent	
Return half	85.48%
Return nothing	14.52%
Number of observations	62

Overall, in all of the nine sessions, the course of action that received most votes was the social optimum, prescribing senders to *send two tokens, and receivers to always share evenly with the sender the sum in their possession.* Hence, in all sessions, this was the course of action that was announced as having attracted the highest consensus.

As can be seen from Table 3 and Figure 2, sending rates and return rates increased more in the Voting condition than in the control condition. In fact, in the control condition return rates were virtually the same in Round 1 and 2. We test for the equality of distribution between first round and second round decisions per treatment, using Wilcoxon sign rank tests. The null hypothesis is that the actions in the first and second round come from the same distribution. The test takes into account that both decisions were made by the same individual. A significant difference exists for both sending and return rates when two tokens are sent in the Voting treatment, but no difference emerges in the Baseline (see Table 5).

interaction is anonymous. Third, randomization of the treatment guarantee that even in presence of a bias due to previous knowledge, this effect is not interacting with the treatment. For a discussion of the role of previous knowledge on the level of trust, see Glaeser et al. (2000), where the interaction is non-anonymous.

Table 5 Wilcoxon signed rank test over hypothesis of equality of distribution between first and second decision

Decision/Condition	Baseline	Voting
Amount sent	z = -0.60 (p = 0.54)	z = -2.37 (p = 0.01)
Return Rate, if 1 token is sent	z = 0.00 (p = 1.00)	z = -1.00 (p =0.31)
Return Rate, if 2 tokens are sent	z = 0.00 (p = 1.00)	z = -2.13 (p = 0.03)

In Table 6, we confirm the same result using a Wilcoxon sign test instead of a sign rank test. The sign test only considers the sign of the difference between the first and the second choice instead of the actual size of the difference. The sign test operates a comparison of the median, whose distribution is binomial (and thus only the p-value is reported). In the Table, we report the relevant test, which is for the one-tailed test over the null hypothesis that the median of the second choice is larger than that of the first one. Again, it can be seen that the results are robust.

Table 6 Signed test over hypothesis that medians of the first decision are lower than medians of the second decision

Decision/Condition	Control	Voting Mechanism
Amount sent	p = 0.39	p = 0.01
Return Rate, if 1 token is sent	p = 0.65	p =0.25
Return Rate, if 2 tokens are sent	p = 0.63	p = 0.02

The treatment effect can also be assessed through a stronger test assessing directly between-treatment differences. We construct a variable given by the difference of behavior in Round 1 and Round 2. This variable will be equal to zero for subjects who have chosen the same action in both rounds. The presence of many zeros harms the power of the test, thus we exclude such cases.

We then run a Fisher exact test on the null hypothesis that the conditional distributions gathered in the Baseline and Voting conditions of the variables so obtained comes from the same distribution. We define the sending rate as the number of tokens being sent divided by two, and the return rate as the number of tokens returned divided by the sum available to the

receiver. In the case of 'Amount Sent', the p value is 0.12, while in the case of the return rate – averaged over the two possible cases - it is 0.06.

In order to assess the robustness of the results, we use a Kruskal-Wallis test to compare the post-treatment behavior across the three villages and across number of comprehension errors. In all cases the hypothesis of equality of distribution is not rejected (Table 7), meaning that neither the village of residence nor the level of comprehension affected decisions.

Table 7 Kruskall Wallis test on hypothesis of equality of distribution across levels of the control variable, in post treatment decisions

Decision/Condition	Village	Comprehension errors
Amount sent	Chi2 (2) = 1.74 (p = .41)	Chi2 (4) = 1.46 (p = .83)
Return Rate, if 1 token is sent	Chi2 (2) = 1.30 (p = .51)	Chi2 (4) = 4.56 (p = .33)
Return Rate, if 2 tokens are sent	Chi2 (2) = .40 (p = .81)	Chi2 (4) = 1.15 (p = .88)

Finally, it has been pointed out that behavior in TGs may be influenced by risk preferences (see section 2.1). We check for whether this is the case in our experiment by assessing the role of risk preferences in investment and return. Risk attitudes are elicited through a self-reported general assessment of one's own risk attitudes, which was experimentally validated by Dohmen et al. (2005)[8]. We use again a Kruskal-Wallis test. The equality of population is not rejected across level of measured risk propensity for amount sent (chi2(4) = 1.87; p = 0.75); return if one tokens is sent (chi2(4) = 3.31; p = 0.50); return if two tokens are sent (chi2(4) = 1.711; p = 0.78). The same result would hold using a self-reported assessment of attitudes towards financial risk.

4.3 Analysis of the relationship between voting and decisions

As we discussed in section 2.2, a variety of mechanisms can account for the observed increase of pro-sociality in Round 2 of the Voting treatment. Our experimental design prevents us from fully disentangling among them. We can nonetheless carry out a speculative analysis to separate commitment effects and normative expectation effects, in analogy with Vanberg (2008) (section 2.2.1). In our experiment, the commitment effect may be deemed to reinforce the propensity for an individual who voted for a certain plan of action to comply with it. The normative expectations effect would induce people to comply with the most voted norm, because this is the most appropriate behavior according to the majority. This may be done analyzing the relationship between voting behavior and Round 2 decisions in

[8]The text of the relevant question was: "Normally, are you fully prepared to take risks or do you try to avoid them?" The answer was elicited on a five points Likert scale. See Question 4 in the questionnaire reported in the Appendix.

the Voting treatment[9]. We mainly focus on Amount Sent because the existence of more possible options than for Return Rates permits a finer disentanglement of types. Table 8 reports the contingency table for Voting and Amount Sent in Round 2.

Table 8: Distribution of Voting and Amount Sent in Round 2 (Voting treatment)

		Amount Sent Round 2			
		0	1	2	
Vote of "Appropriate" Amount Sent	0	0	0.05	0.05	0.10
	1	0.03	0.24	0.15	0.42
	2	0	0.05	0.44	0.48
		0.03	0.34	0.63	1

We aim at classifying subjects into three mutually exclusive types: commitment-concerned individuals, expectations-concerned individuals, and a residual category including behaviors that cannot be accounted for by the previous two types. We first look at the group who voted that the appropriate course of action for Amount Sent was either zero or one. 52% of participants did so. For these individuals, we can provide an unambiguous classification of participants into types. Within this group, 46% of participants who voted less than two (and 24% of the full sample) decided to stick with their vote of sending one token to the counterpart. We can unambiguously classify these individuals as "commitment-concerned", as the normative expectation hypothesis would prescribe to send two tokens in Round 2. We can then classify as "expectation-concerned" individuals the 38% within this group (20% overall) who decided to send two tokens in Round 2 after having voted for less than two tokens. The remaining 15% of participants (8% overall) do not fall in any of these two categories, so we classify them as belonging to the residual group.

A clear-cut assignment into types cannot be carried out if individuals voted for two tokens, because these individuals may have been motivated by both commitment and expectations in sending two tokens in Round 2. We can nonetheless provide some lower and upper bounds. We should note that 90% of individuals who voted for two tokens as the most appropriate course of action did indeed send two tokens in Round 2. Hence, the share of

[9] We thank an anonymous referee for suggesting to investigate the relationship between voting and behavior.

unclassified individuals is equal to 10% in this case. These individuals may perhaps have been motivated by "strategic" considerations. Voting for two tokens – while only sending one - may have been a way to induce others to send two tokens, thus receiving a larger amount when acting as receivers.

An upper (lower) bound for commitment-concerned individuals can be found assuming that all individuals (none) who voted for two tokens as the most appropriate course of action and sent two tokens in Round 2 were motivated by commitment concerns. Computing a weighted average of observed frequencies when Vote is either less than two or equal to two yields that commitment-concerned individuals may range from a minimum of 24% up to a maximum of 67%, while normative-concerned individuals may be quantified as varying from a minimum of 20% up to a maximum of 63%[10]. A conservative assumption may be that individuals are equally split between these two types when they voted for two tokens. This would result in 46% of individuals being classified as "commitment-concerned", 41% as "expectation-concerned", while 13% would not be classified.

Clearly, these calculations can only be considered as "back-of-the-envelope" and do not have pretense of generalizability. Moreover, they were carried out under the assumption that concerns with commitment and expectations were the only main factors of behavior. Other factors, such as a desire to conform with the most voted norm, or the realization of the intrinsic value of the norm after voting, may have also have been acting as motivators. Individuals who voted for one token and sent one token after the most voted norm was announced may have simply believed that sending one token was the optimal action regardless of any desire to commit with their vote. Obviously, neglecting these additional

[10] For instance, to compute the upper and lower bounds of commitment-concerned individuals, we consider the following formula for conditional probabilities: $rob(Type = C) = Prob(Type = C|Vote < 2)\ X\ Prob(Vote < 2) +\ Prob(Type = C|Vote = 2)\ X\ Prob(Vote = 2)$.
$Prob(Type = C)$ is the probability that an individual is commitment-concerned. Assuming that probabilities coincide with observed frequencies, we can then quantify probabilities as follows: $Prob(Type = C|Vote < 2) = 0.46$; $Prob(Vote < 2) = 0.52$; $Prob(Vote = 2) = 0.48$. As we cannot quantify $Prob(Type = C|Vote = 2)$, we can compute lower and upper bounds by observing that this probability can range from a minimum of 0 up to a maximum of 0.90. As observed, 10% of individuals are neither commitment-concerned nor expectation-concerned when Vote=2, hence the maximum is 0.90 rather than 1. Replacing these probabilities into the above formula yields the estimates reported in the text. Similar formulas can be used to compute the lower and upper bounds for expectation-concerned types.

factors, as we did, lead to an over-estimation of the estimated size of the two types under consideration. In spite of all these caveats, we tentatively conclude that both motivational factors – that is, commitment and expectations – appear to have been relevant in determining individuals' behavior in Round 2, with the commitment effect being marginally dominant over the expectation effect.

Even if we cannot exactly pinpoint the individual contribution of either commitment or expectations on behavior, we can nonetheless appreciate their overall aggregate impact. Firstly, a chi squared test soundly rejects the null hypothesis of independence of the distribution of Voting and Amount Sent in Round 2 ($chi2(4) = 19.60$, $p = 0.001$). This proves that voting clearly affected behavior in Round 2, as far as Amount Sent is concerned. Secondly, we compare the distribution of choices for Amount Sent in Round 1 and Round in the two experimental conditions in Figure 3. It is striking that a much larger percentage of subjects (47%) sent two tokens in both rounds of the Voting condition than in the Baseline (21%). Given the similarity of behavior in the first round between Baseline and Voting conditions, this result must be due to voting inducing many individuals to stick with sending two tokens rather than switching to another strategy. As many as 37% of subjects participating in the Voting treatment sent two tokens in both Round 1 and 2 and voted two tokens. 10% sent two tokens in both rounds but voted for less than two. We can conjecture that, without voting, less than half of people who sent two tokens in Round 1 would have done the same in Round 2. The introduction of voting made many more people willing to send two tokens in both rounds than what it would have been the case without voting.

Figure 3: Relationship between Round 1 and Round 2 Amount Sent per treatment condition

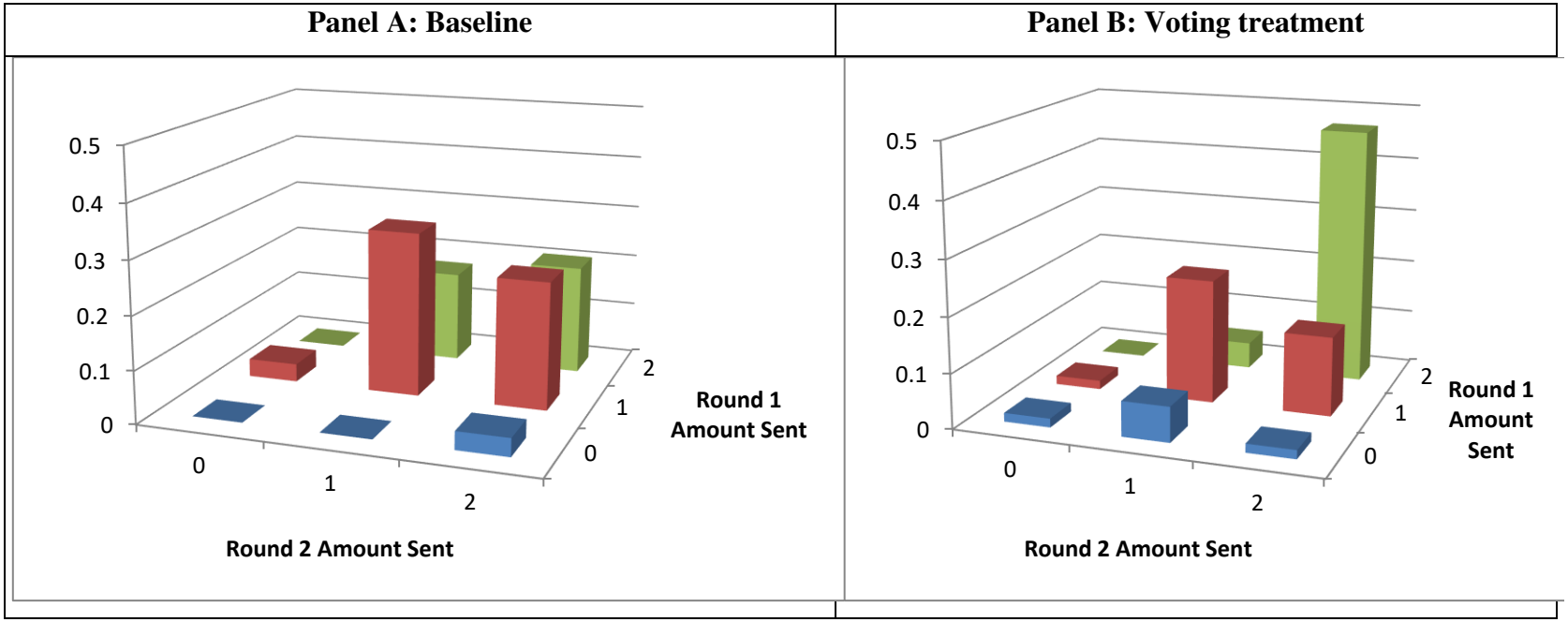


Figure 4: Relationship between Round 1 and Round 2 Return Rate (if Amount Sent =1) per treatment condition

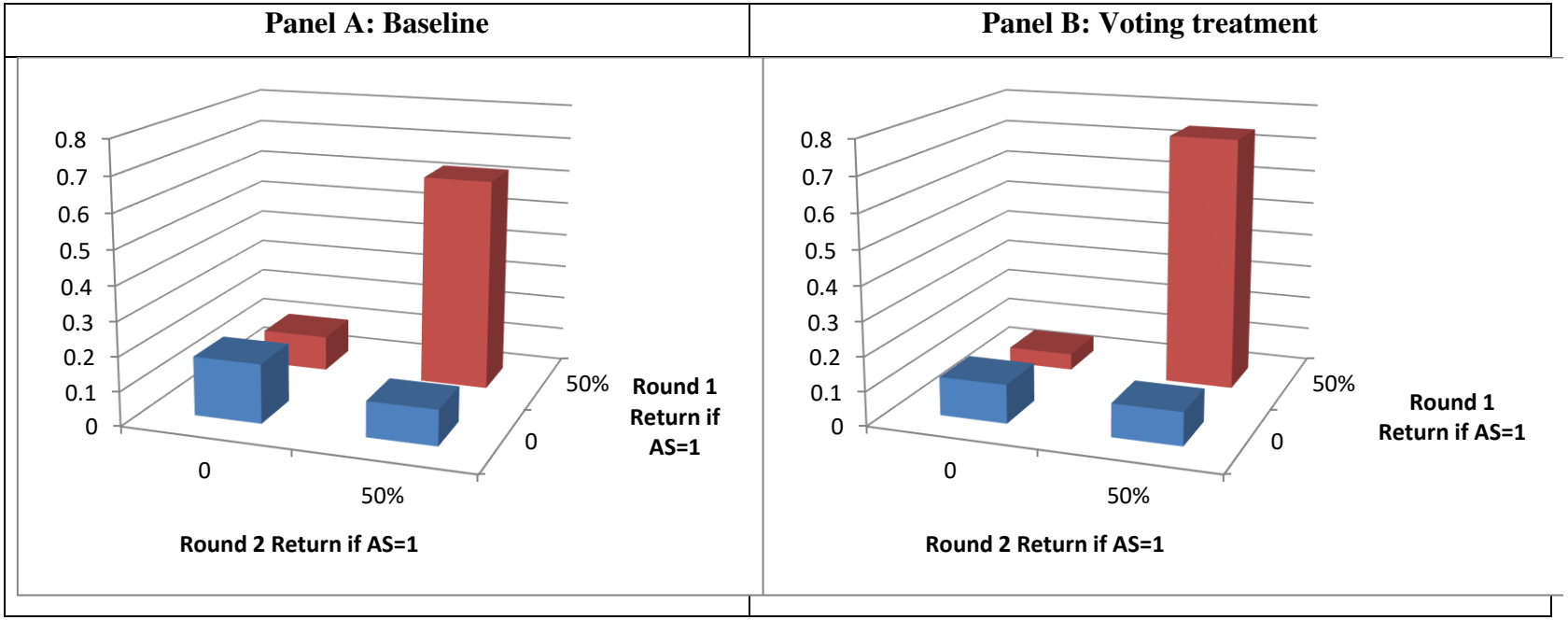


Figure 5: Relationship between Round 1 and Round 2 Return Rate (if Amount Sent =2) per treatment condition

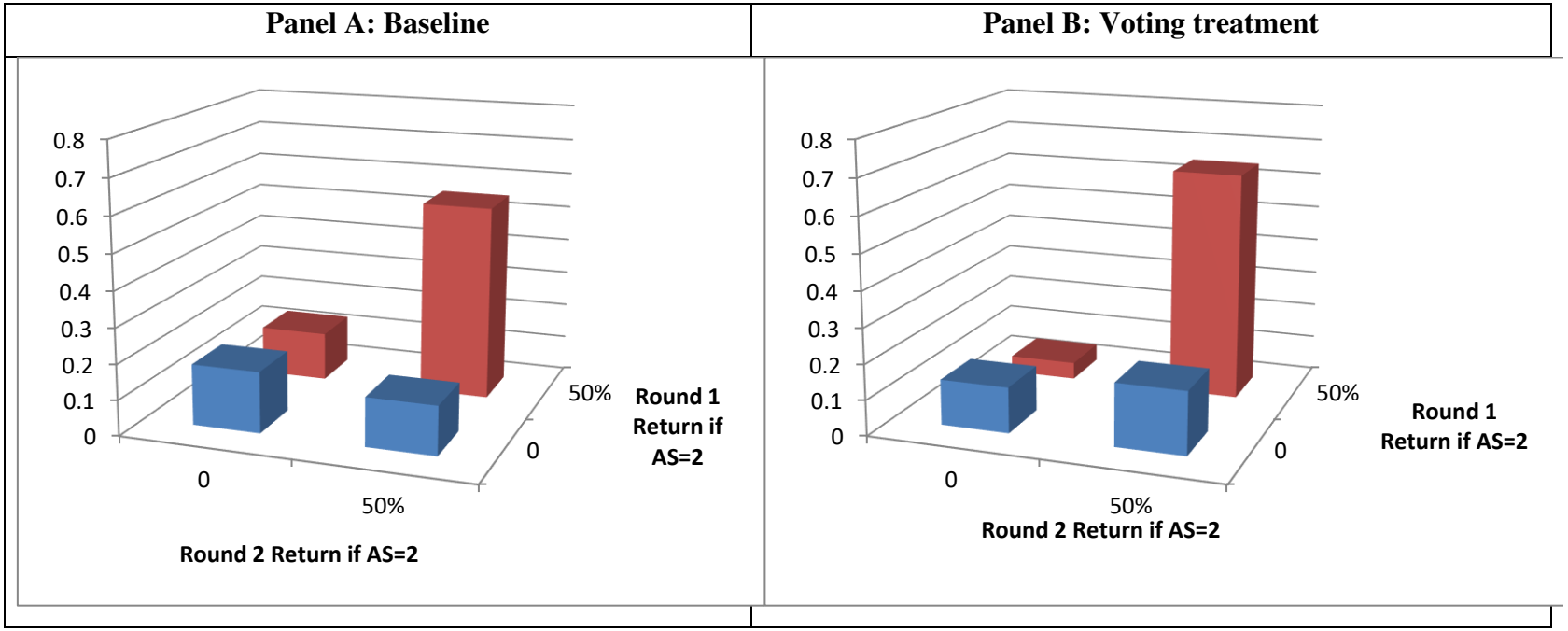


Table 9: Changes of Amount Sent in Round 1 and 2 per experimental condition

	Baseline	Voting treatment
Percentage of subjects for whom $AS_1 > AS_2$	21%	6%
Percentage of subjects for whom $AS_1 = AS_2$	52%	71%
Percentage of subjects for whom $AS_1 < AS_2$	28%	23%

It is also interesting to note that the number of subjects who exactly replicated their choices in Round 1 and 2 is considerably lower in the Baseline (52%) than in the Voting condition (71%) (Table 9). Clearly, subjects in the Baseline condition did not see the two Rounds as independent. Arguably, they operated some kind of diversification in their portfolio choice, possibly as an insurance mechanism device, in spite of only one decision being determinant for payoffs. Interestingly, while the number of subjects who increased or decreased their Amount Sent roughly canceled each other out in the Baseline, the fraction of subjects increasing their contribution from Round 1 to Round 2 appears significantly higher in the Voting treatment. Statistical tests bear out this impression. In the Baseline condition, a Binomial test does not reject the null hypothesis that the number of subjects who increased their contribution is equal to those who reduced it (p=0.39; n=14), while it does reject the null in the Voting condition (p=0.02, n=18). Similarly, a chi squared test does not reject the null hypothesis of independence of the distributions of Amount Sent in Round 1 and 2 in Baseline (chi2(4) = 2.10, p = 0.72), while it does so in the Voting treatment (chi2(4) = 24.68, p< 0.001).

We report the distribution of choices for return rates in Figures 4 and 5. We note a general tendency for subjects to choose the fair split in both rounds in the Voting treatment and in the Baseline, for either case of AS=1 and AS=2. Nevertheless, differences between distributions seem less pronounced than for Amount Sent because of the overall tendency to prefer the fair split to the alternative in both conditions, and because of the lower number of alternatives. Nonetheless, we note that chi squared independence tests tend to reject the null

hypothesis more strongly in the Voting treatment than in the Baseline condition[11]. We conjecture that this difference is due to the tendency for return rates to increase over rounds in the Voting treatment, while they tend to remain constant in the Baseline, as we observed for Amount Sent. In other words, the number of people switching from the fair to the unfair split between rounds canceled each other out in Baseline, while subjects tended to increase their return rates in the Voting treatment.

As for the relationship between Vote and Round 2 return rate, the fact that too many choices clustered on voting for the fair split and complying with it makes it pointless to replicate the same exercise carried out above for the quantification of types with respect to Amount Sent. Nonetheless, we note an interesting result. The Chi square of independence of Vote and Round 2 Return Rate is not rejected for AS=1 (chi2(1) = 2.3036, p= 0.13), but is rejected at strong significance levels for AS=2 (chi2(1) = 10.32, p = 0.001). This suggests that people's chosen return rates adhered more to their previous vote for the norm that was announced as the most voted one, than for the other norm.

5. Discussion

Results clearly indicate an increase in pro-social behavior- specifically, an increase of first-movers' transfers and of second-movers' return rates when AS=2 – in the Voting treatment, while no significant effect is detected in the Baseline. Given the speculative character of our analysis in section 4.3 to classify subjects into behavioral types, the main limitation of our study is probably the incapacity to pin down the underlying mechanisms for our results. In particular, it would have been interesting to separate the effect of the voting mechanism from the effect of the announcement of a course of action, which may have made a norm of reciprocity salient. However, this would have required a sample size far exceeding that available in this experiment. An alternative route may have been to elicit beliefs of first

[11] As far as Return Rate for AS=1 are concerned, the null hypothesis of independence between Round 1 and Round 2 decisions is rejected at strong significance levels in both Baseline (chi2(1) = 6.74, p = 0.009) and Voting treatment (chi2(1) = 17.30, p<0.001). As for Return Rate for AS=2, the null hypothesis of independence between Round 1 and Round 2 decisions is only weakly rejected at 10% significance in Baseline (chi2(1) = 3.67, p = 0.056), while it is strongly rejected in the Voting treatment (chi2(1) = 11.14, p=0.001).

and perhaps second order on others' preferences and others' voting behavior. We preferred not to do this for the danger that belief elicitation may influence future choices – if conducted beforehand (Rutström and Wilcox, 2009; Gächter and Renner, 2010; Blanco et al., 2010) – or being mere ex-post rationalization of observed behavior – if conducted afterwards (Nickerson, 1998). We believe that establishing a link between voting and pro-sociality in the cleanest possible way, without the risk that the belief decision may somehow "contaminate" the choice decision, was an important enough objective to achieve.

The fact that our results are not merely due to an announcement effect, but are at least partly due to voting, can perhaps be best borne out comparing our results with the seminal paper by Berg et al. (1995). As already mentioned in section 2.1, while in their experiments no impact from announcing previous behavior was found on first-movers' transfers, in our experiment we do find such an effect. This shows, albeit qualitatively, that the effect of voting goes beyond the mere effect of announcing a norm of behavior in a TG. Moreover, no social norm seems to regulate, in general terms, first-movers' behavior. Return rates, conversely, are subject to an easily identifiable social norm prescribing the repayment of the trustor's transfer (Bicchieri et al., 2011). It is therefore not so surprising that the voting mechanism influenced behavior that, within a TG, is less likely to be driven by social norms. Moreover, experiments testing the impact of institutions on pro-social behavior generally find that the introduction of an institution as a result of a vote is considerably higher than the introduction of the same institution without a vote (Tyran and Feld, 2006; Rauchdobler et al., 2010).

Some clarifications are necessary with regards to external validity. Following the approach of Campbell and Cronbach (Shadish et al. 2002; Cronbach et al. 1980), inference from local experimental evidence should be analyzed as the possibility to make inference from specific UTOS (unit, treatment, observation and settings) to general ones (UTOS* following Cronbach's notation); this is of course a more detailed discussion, which goes beyond representativeness of the sample (i.e. units). Starting from the latter, the sample has not been designed statistically, and as a result, is not representative of the population of the three villages. The treatment (the Voting condition) is designed and explained following standard practice of consultation (secret ballot with majority voting) and certainly meets the requirement of external and construct validity. The same could be said about the setting: the

use of a lab in the field in the three sites increases the likelihood of correct inference. Finally, regarding the observation, i.e. the concrete measurement performed, there is no consensus on how to measure trust. Part of the literature, especially in social sciences other than economics, tend to use general questions, such as that of World Value Survey (e.g. "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?" or similar phrasing). According to Fehr (2009), this question has the merit of capturing both preferences and beliefs, which affect trust behavior. However, this question is not preferable to a behavioral measure of trust, such as the one elicited in the present study, because it does not capture the act of transferring resources to the availability of other parties, which is the peculiar characteristic of trust. Moreover, an answer to the survey question is the result of introspection of one's behavior in a real-life setting, and as such, it leaves many aspects open to subjective interpretation. A behavioral measure increases the precision of the elicited response because all the contingencies of one's action are fully spelled out (Fehr, 2009).

We do not see as problematic that data comes from three villages, because there is cultural homogeneity across villages. We did not have participants from ethnic minorities, which are very rare in these locations. Most importantly, we showed that behavior is not systematically different across sites (Table 7).

We conclude this discussion with some considerations on whether the possible deviations in behavior in the second block of the experiment may be due to an experimenter demand effect (Zizzo, 2010). Experimenter demand effects may be said to occur when participants in the experiment are induced to act in a certain way because they perceive that a course of action is more or less explicitly suggested by the experimenter. Zizzo (2010) distinguishes between social and cognitive experimenter demand effects. The former are driven by observations of others' behavior or by the processing of cues associated with other subjects' behavior. The latter are due to changes in the construal of the interaction that the subject can carry out as an effect of cues included in the instructions or in the procedures of the experiment. Examples of cognitive experimenter demand effects include the explicit mention in the instructions that participants should maximize their payoffs, or the

recommendation that players should take care of other players when making distribution decisions in Dictator Games.

It is clear that in our experimental design the communication of the voting procedure may be taken to represent a form of social pressure, thus making the possibility of social experimenter demand effects concrete. We nonetheless believe that this risk is on the one hand limited and on the other hand implicit in the very situation we want to study. First, the publicity of the information referred to the result of a vote, rather than to actual behavior. The risk of social pressure is therefore less strong than what it would have been otherwise. Moreover, it was clearly stated that the communication of the vote was completely non-binding. When deciding how much to send and how much to return, subjects were risking or sacrificing a substantial amount of their own money. We find it hard to believe that they did so exclusively "to please" the experimenter. Conversely, we find plausible that subjects recognized that the sharing of the information regarding the vote made the existence of social norms of reciprocity salient in the group. This made trustors more willing to risk their money and trustees more eager to reciprocate the observed behavior with increased return rates. We would also like to point out that, in order to characterize correctly consultative democracy, the result of the voting should be made public. It is then inevitable that social pressure played a part in our experiment, for the same reasons it would play a role in real life. Finally, it should be stressed that the threat of cognitive experimenter demand effects seems minimal in our setting. The language used in the instructions was neutral and we refrained from using possibly loaded terms, such as "democracy", "collective", or "group" – insofar as the voting procedure was concerned. We conclude that undoubtedly further research should ascertain the extent to which the observed results were due to subjects' perception of the experimenters' desired behavior. Based on the above considerations, we believe that this risk is on the one hand instrumental to the study of the research question, and on the other hand, the monetary incentives were such that this effect should have had minimal consequences on subjects' behavior.

6. Conclusions

In this article, we have investigated the role of a consultation with democratic rules as a mechanism to promote pro-social behavior in a TG. We have shown that the introduction of such a consultation significantly increases both trust and trustworthiness. Results are robust to the consideration of village effects, comprehension, and risk preferences.

We have discussed that the main limitation of our study is the inability to shed light on the mechanisms underlying our results. Even if we cannot fully explore the *motivational* level – that is, the individual motivations that triggered our results – we believe that at the *institutional* level our results are clear-cut. The introduction of a voting procedure in our experiment, which shares obvious resemblance with a process of consultative democracy in real-life, shifts social behavior closer to the social optimum. This result resonates with the finding that democratic institutions are associated with higher levels of trust (Ljunge, 2014).

We believe that our results have potentially wide-ranging policy implications, although further tests on its generalizability to other types of social interactions - e.g. simultaneous cooperation games – should be carried out. The main policy implication is that public consultation regarding the appropriate behavior that people should hold in social interactions seems to be effective in promoting pro-social behavior. This calls for implementing public consultations for many aspects of collective decisions, not only for political representation. The execution of public projects, the allocation of public funds to different welfare or health programs, the indication of priorities in the administration of resources, are all cases in which citizens may be consulted in order to express their "voice" and shape public governance. Clearly the use of consultative democracy should be limited, both to avoid "saturation effects" from the use of the instrument, and not to pre-empt the activities of public administration. Nevertheless, the scope for greater adoption of this instrument seems considerable.

It should be stressed that it is important that the results of the consultation be released to the population. In fact, a standard "nudging" technique consists of disseminating information on the average behavior in a population over relevant issues. This has normally the effect of spurring people below the average to act in a way closer to the announced average behavior. Admittedly, it may also have the same effect – albeit with a contrary "sign" – for people above the norm. For instance, these people may be prompted to decrease their

levels of cooperation when informed of the population's average behavior. Overall, the extant evidence seems to suggest that the positive effect of raising the "under-performers" may more than offset the negative effect on "over-performers" (Frey and Meier, 2004; Shang and Croson, 2009; Agerström et al., 2016). Arguably, this may be particularly the case if people are asked to vote on the "most appropriate" course of action, rather than on their customary course of action. Concerns for fairness and justice would probably determine a majority for a norm well above the average standard of behavior. Overall, we are inclined to think that the combination of the "voting" effect and of the announcement effect may be overall beneficial to the community.

The fact that our results were obtained from samples of predominantly low-educated, low-income people from rural communities, makes them particularly relevant for policy. These are the people who would be the recipients of public programs and who play a key role to solve the problems of underdevelopment and poverty traps that affect these communities. It is therefore important to focus on them.

References

Agerström, J., Carlsson, R., Nicklasson, L., & Guntell, L. (2016). Using descriptive social norms to increase charitable giving: The power of local norms. *Journal of Economic Psychology*, 52, 147-153.

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, *115*(3), 715-753.

Anand, P. (2001). Procedural fairness in economic and social choice: Evidence from a survey of voters. *Journal of Economic Psychology*, 22*(2)*, 247-270.

Arrow, K. (1974). *The limits of organization.* New York: Norton.

Bacharach, M. (1999). Interactive team reasoning: a contribution to the theory of co-operation. Research in economics, 53(2), 117-147.

Ball, A., Wolf, C., Ocklenburg, S., Herrmann, B., Pinnow, M., Brüne, M., et al. (2013). Variability in ratings of trustworthiness across the menstrual cycle. *Biological psychology*, 93, 52-57.

Balliet, D. (2010). Communication and cooperation in social dilemmas: A meta-analytic review. Journal of Conflict Resolution, 54(1), 39-57.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. Journal of personality and social psychology, 71(2), 230.

Battigalli, P., & Dufwenberg, M. (2007). Guilt in games. The American economic review, 97(2), 170-176.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychological bulletin, 117(3), 497.

Becker, A. & Miller, L.(2009). Promoting justice by treating people unequally: an experimental study. *Experimental Economics,* 12*(4),* 437-449.

Belli, S., Rogers, R. & Lau, J. (2012). Adult and adolescent social reciprocity: Experimental data from the Trust Game. *Journal of Adolescence*, 35, 1341-1349.

Berg, J., Dickhaut, J. & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10, 122-142.

Bicchieri, C & Xiao, E. (2009). Do the Right Thing: But Only If Others Do So. *Journal of Behavioral Decision Making*, 22, 191-208.

Bicchieri, C., Xiao, E., & Muldoon, R. (2011). Trustworthiness is a social norm, but trusting is not. Politics, Philosophy & Economics, 10(2), 170-187.

Bidner, C and P Francois (2011) Cultivating Trust: Norms, Institutions, and the Implication of Scale. *Economic Journal,* 121(555): 1097-1129

Blanco, M., Engelmann, D., Koch, A. K., & Normann, H. T. (2010). Belief elicitation in experiments: is there a hedging problem?. *Experimental Economics*, *13*(4), 412-438.

Boero, R., Bravo, G., Castellani, M. & Squazzoni, F. (2009). Reputational cues in repeated games. *The Journal of Socio-Economics*, 38, 871-877.

Bogliacino, F and P Ortoleva (2015) The Behavior of Other as a Reference Point, DOCUMENTOS DE TRABAJO - ESCUELA DE ECONOMÍA 013611, UN - RCE - CID.

Bohnet, I., Greig, F., Herrmann , B. and Richard Zeckhauser (2008). Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and theUnited States, *American Economic Review*, 98 (1), 294-310.

Bolton, G. E., Brandts, J., & Ockenfels, A. (2005). Fair procedures: Evidence from games involving lotteries. *Economic Journal*, 115(506), 1054-1076.

Bonilla, J., López, D. & Sepúlveda, C. (2014). Chapter 2. Estratificación socioeconómica en Colombia. Contexto general y diagnóstico. In: *Los límites de la estratificación: en busca de alternativas,* eds. C.E. Sepúlveda, D. López and J.M. Gallego. Bogotá D.C.: Editorial Universidad del Rosario.

Bracht, J. & Feltovich, N. (2009).Whatever you say, your reputation precedes you: observation and cheap talk in the trust game. *Journal of Public Economics*, 93*(9-10)*, 1036-1044.

Buchan, N., Croson, R. & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the Investment Game. *Journal of Economic Behavior & Organization*, 68, 466-476.

Camerer, C. F. (2003). *Behavioral Game Theory*. Princeton: Princeton University Press.

Cárdenas, J. C. & Chong, A. & Ñopo, H.(2013) Stated social behavior and revealed actions: Evidence from six Latin American countries, Journal of Development Economics, 104(C), p. 16-33

Casari, M. & Cason, T.(2009). The strategy method lowers measured trust worthy behavior. *Economics Letters*, 103, 157-159.

Charness, G., & Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, *74*(6), 1579-1601.

Charness, G., Cobo-Reyes, R. & Jiménez, N. (2008). An investment game with third-party intervention. *Journal of Economic Behavior & Organization*, 68, 18-28.

Charness, G., Du, N. & Yang, C. (2011). Trust and trustworthiness reputations in an investment game. *Games and Economic Behavior*, 72, 361-375.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of personality and social psychology*, 58(6), 1015.

Cialdini, R.B. (1993). Influence: Science and Practice, 3rd edition, New York.

Cochard, F., Van, P. & Willinger, M. (2004). Trusting behavior in a repeated investment game. *Journal of Economic Behavior & Organization*, 55, 31-44.

Coleman, J. (1990). *Foundations of Social Theory*. Cambridge: Harvard University Press.

Cox, J. C. (2004). How to identify trust and reciprocity. *Games and economic behavior*, 46(2), 260-281.

Cronbach, LJ, S.R. Ambron, S.M. Dornbusch, R.D. Hess, R.C. Hornik , D.C. Phillips, D.E. Walker, SS Weiner: Toward reform of program evaluation, San Francisco 1980, Jossey-Bass.

Dawes, R. M. (1980). Social dilemmas. Annual review of psychology, 31(1), 169-193.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. & Wagner, G. (2005).Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey. DIW Berlin, Discussion Paper511.

Dubois, D., Willinger, M. & Blayac, T. (2012). Does players' identification affect trust and reciprocity in the lab? *Journal of Economic Psychology*, 33, 303-317.

Duffy, J., Xie, H. & Lee, Y-J. (2013).Social norms, information, and trust among strangers: theory and evidence. *Economic Theory*, 52*(2)*, 669-708.

Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and economic behavior*, *47*(2), 268-298.

Eckel, C. C., & Wilson, R. K. (2004). Is trust a risky decision?. *Journal of Economic Behavior & Organization*, 55(4), 447-465.

Ellingsen, T and M Johannesson (2004) Promises, Threats and Fairness. *Economic Journal,* 114(495): 397-420

Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and economic behavior*, 54(2), 293-315.

Fehr, E. (2009). On the economics and biology of trust. *Journal of the European Economic Association*, 7, 235-266.

Fehr, E. and Gächter, S. (2002) Altruistic punishment in humans. Nature 415, 137–140

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and human behavior*, *25*(2), 63-87.

Fehr, E., and Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*, 422(6928), 137-140.

Franzen, N., Hagenhoff, M., Baer, N., Schmidt, A., Mier, D., Sammer, G., Gallhofer, B., Kirsch, P.& Lis, S. (2011). Superior 'theory of mind' in borderline personality disorder: An analysis of interaction behavior in a virtual trust game. *Psychiatry Research*, 187, 224-233.

Frey, B. & Stutzer, A. (2005). Beyond outcomes: measuring procedural utility. *Oxford Economic Papers,* 57, 90-111.

Frey, Bruno, S., and Stephan Meier. 2004. "Social Comparisons and Pro-social Behavior: Testing "Conditional Cooperation" in a Field Experiment." *American Economic Review,* 94(5): 1717-1722.

Friedman, D. and Sunder, S. (1994). Experimental Methods: A Primer for Economists. Cambridge: Cambridge University Press.

Fudenberg, D., & Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 533-554.

Gächter, S., & Renner, E. (2010). The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics*, *13*(3), 364-377.

Gächter, S., Herrmann, B., & Thöni, C. (2010). Culture and cooperation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *365*(1553), 2651-2661.

Glaeser, EL, DI Laibson, JA Sheinkman, and CL Soutter (2000) Measuring Trust. *Quarterly Journal of Economics,* 115(3): 811-846

Grimalda, G., Kar, A., & Proto, E. (2016). Procedural fairness in lotteries assigning initial roles in a dynamic setting. Experimental Economics, 19(4), 819-841.

Guiso, L., Sapienza, P. & Zingales, L.(2006). Does culture affects economic outcomes? *Journal of Economic Perspectives*, 20*(2)*, 23-48.

Hardin, R. (2002). *Trust and trustworthiness*. Russell Sage Foundation.

Harrison, GW, List, JA (2004) Field Experiments. Journal of Economic Literature, 42(4): 1009-1055

Henrich, J, Heine, SJ, Norenzayan, A. (2010) The weirdest people in the world? BEHAVIORAL AND BRAIN SCIENCES (2010), 33: 1-75

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, *319*(5868), 1362-1367.

Houser, D., Schunk, D. & Winter, J. (2010). Distinguishing trust from risk: An anatomy of the investment game. *Journal of Economic Behavior & Organization*, 74, 72-81.

Houser, D., Schunk, D., & Winter, J. (2010). Distinguishing trust from risk: An anatomy of the investment game. *Journal of economic behavior and organization*, 74(1), 72-81.

Karapetyan, D. & d'Adda, G. (2014). Determinants of conservation among the rural poor: A charitable contribution experiment. *Ecological Economics*, 99, 74-87.

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

Karni, E. & Safra, Z. (2002). Individual sense of justice: a utility representation. *Econometrica*, 70, 263-284.

Karni, E., Salmon, T. &Sopher, B. (2008). Individual sense of fairness: an experimental study. *Experimental Economics*, 11*(2)*, 174-189.

Klackl, J., Pfundmair, M., Agroskin, D. & Jonas, E. (2013). Who is to blame? Oxytocin promotes non personalistic attributions in response to a trust betrayal. *Biological Psychology*, 92, 387-394.

Kleine, F., Königstein, M. & Rozsnyói, B. (2014). Voluntary leadership in an experimental trust game. *Journal of Economic Behavior & Organization* [Forthcoming]

Knack, S. & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics*, 112, 1251-1288.

Kollock, P. (1994). The emergence of exchange structures: An experimental study of uncertainty, commitment and trust. *American Sociology Review*, 100, 313-345.

Korsgaard, M. A., Schweiger, D. M., & Sapienza, H. J. (1995). Building commitment, attachment, and trust in strategic decision-making teams: The role of procedural justice. Academy of Management journal, 38(1), 60-84.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. The Quarterly Journal of Economics, 121(4), 1133-1165.

Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of personality and social psychology*, *46*(5), 1044.

Krawczyk, M. (2011). A model of procedural and distributive fairness. *Theory and Decision*, 70*(1)*, 111-128.

Krawczyk, M.&Le Lec, F. (2010). Give me a chance! An experiment in social decision under risk. *Journal of Experimental Economics*, 13, 500-511.

La Porta, R., López-de-Silanes, F., Shleifer, A. & Wishny, R. (1999). The quality of government. Journal of Law, Economics and Organization, 15(1), 222-279.

Lei, V. & Vesely, F. (2010). In-Group versus Out-Group Trust: The Impact of Income Inequality. *Southern Economic Journal*, 76*(4)*, 1049-1063.

Lind, E. & Tyler, T. (1988). *The social psychology of procedural justice*. Springer.

Ljunge, M. (2014) Social capital and political institutions: Evidence that democracy fosters trust. *Economics Letters,* 122(1): 44-49

Lunawat, R. (2013). An experimental investigation of reputation effects of disclosure in an investment/trust game. *Journal of Economic Behavior & Organization*, 94, 130-144.

McClennen, E. F. (1990). *Rationality and dynamic choice: Foundational explorations*. Cambridge university press.

Nederhof A. Methods for coping with social desirability bias: a review. European Journal of Social Psychology. 1985; 15: 263–280.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, *2*(2), 175.

Olken, B. A. (2010). Direct democracy and local public goods: Evidence from a field experiment in Indonesia. *American Political Science Review*, 104*(02)*, 243-267.

Ostrom, E., & Walker, J. (1991). Communication in a commons: cooperation without external enforcement. Laboratory research in political economy, 287-322.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American political science Review*, *86*(2), 404-417.

Palhaus D. Two-component models of Socially Desirable Responding. Journal of Personality and Social Psychology. 1984; 46(3): 598–609.

Putnam, R., Leonardi, R. & Nannetti, R.(1993). *Making democracy work: civic traditions in modern Italy.* Princeton: Princeton University Press.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American economic review*, 1281-1302.

Rauchdobler, J., Sausgruber, R, Tyran JR (2010) Voting on Thresholds for Public Goods: Experimental Evidence, *FinanzArchiv: Public Finance Analysis*, 66(1), 34-64

Rietz, T., Sheremata, R., Shields, T. & Smith, V. (2013). Transparency, efficiency and the distribution of economic welfare in pass-through investment trust games. *Journal of Economic Behavior & Organization*, 94, 257-267.

Rutström, E. E., & Wilcox, N. T. (2009). Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test. *Games and Economic Behavior*, *67*(2), 616-632.

Sampson, R., Raudenbush, S. & Earls, F. (1997). Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*, 277, 918-924.

Selten, R., Abbink, K., Buchta, J. & Sadrieh, A. (2003) How to Play 3x3-Games A Strategy MethodExperiment. *Games and Economic Behavior*, 45*(1)*, 19-37.

Shadish, WR, T.D. Cook, D.T. Campbell: Experimental and Quasi-Experimental Designs for Generalized Causal Inference, Boston, MA 2002, Houghton Miffl in Company

Shang, J., Croson, R., 2009. A field eexperiment in charitable contribution: The impact of social information on the voluntary provision of public goods. Economic Journal, 119, 1422–1439.

Smith, V. (1976). Experimental Economics: Induced Value Theory. American Economic Review, 66(2), 274-279.

Sugden, R. (1986). The economics of rights, co-operation and welfare (Vol. 22). Oxford: Basil Blackwell.

Thibaut, J. W., & Walker, L. (1975). *Procedural justice: A psychological analysis*. New York: L. Erlbaum Associates. pp. 1141-1160.

Trautmann, S. (2009). A tractable model of process fairness under risk. *Journal of Economic Psychology*, 30*(5)*, 803-813.

Tyler, T. (2006). *Why people obey the law*. Princeton University Press.

Tyran, J.-R. and Feld, L. (2006): Achieving Compliance when Legal Sanctions are Non-Deterrent. *Scandinavian Journal of Economics* 108(1): 135-56

Vanberg, C (2008) Why Do People Keep Their Promises? An Experimental Test of Two Explanations. *Econometrica*, 76(6): 1467-1480

Zak, P. J., & Knack, S. (2001). Trust and growth. *Economic Journal*, 111(470), 295-321.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75-98.

Appendix: Supplementary Materials

Experimental protocol[12]

Participants were recruited through various channels (including word of mouth and presentations during community meetings in the village hall and the local church). We introduced ourselves as University researchers conducting a research project, for which we were recruiting participants in incentivized decision-making tasks. Subjects had to sign an informed consent to participate.

All the sessions took place during weekends. We run three sessions in a single day in each village. This minimized contagion or collusion effects. In all the three locations we guaranteed confidentiality and anonymity ensuring sufficient distance among the tables and/or the use of separators (boxes). The experiments were conducted in November 2013 (Junin, Cota) and January 2014 (La Mesa).

All the session were led by a fluent Spanish speaker, with the help of assistants who would hand out materials, entry individuals' decisions in the computer, and provide assistance to subjects with literacy problems.

When the participants arrive at the session they enter the first room one by one. They receive the informed consent form and should return signed before proceeding. At their entrance, they have to draw an envelope from either an even-numbered ID deck or an odd-numbered ID deck. The envelopes are identical except for the code appearing on them, which is not shown to the participants. Previously we had randomly drawn whether the odd-numbered ID deck or the even-numbered ID-deck would be shownfirst to participants. For example, if we draw "even", the first participant has to select among the even-numbered ID deck, then the second among the odd-numbered ID deck, etcetera.This ensures an equal number of senders and receivers in the TG. We explain that the code is important to guarantee participants' anonymity throughout the study. We report the ID numbers that are selected into an Excel spreadsheet and collect the informed consent form. Participants can only show their ID-number to the researchers. We explain that the ID number is required only for their

[12] Instructions that should not be read are in italic.

identification when final payments are handed out. Finally participants can enter the experimental room, sitting at the sector that is associated with the "even" or "odd" ID-numbers.

Welcome. We thank you for your participation in this exercise, as part of a research about individual decision-making. This research is coordinated by the Fundación Universitaria Konrad Lorenz, from Bogotá. This session will last approximately two hours.

Depending on the decisions made by you and other participants, you may win an important amount of money, and that is why you should pay attention to these instructions. All the information given by you to this research is absolutely confidential and will be used only for academic purposes. Your name or ID won't be shown in any report of this research.

It is important for you to know that you can leave at any time. However, you can receive the final payment only if you take party to all the decisions and fill in the final questionnaire.

Now I am going to tell you how the final payments are determined: these payments are the result of the decisions you make over four activities. At the end you will receive the payment of only one of these activities. The activity that is going to be paid will be chosen randomly in front of you once everybody completes all the activities. It is important that you think carefully at your decisions on each one of the activities, because you will only know which one of the activities is going to be paid at the very end. Once you finish all the activities and the activity among the four that is paid is chosen, we will pass you a questionnaire while we calculate your payment. We will distribute the payments at the end of the session.

GROUP INSTRUCTIONS FOR THE FIRST CHOICE

1) Each one of you will interact with another person. You will not know the identity of your counterpart. How much money you will earn depends on the decisions made by you and you rcounterpart. The final payment will be delivered to you privately at the end of the session.
2) We call the two persons Sender and Receiver. Both of them are inside this room. Before we start, we will tell you whether you will be the Sender or the Receiver.

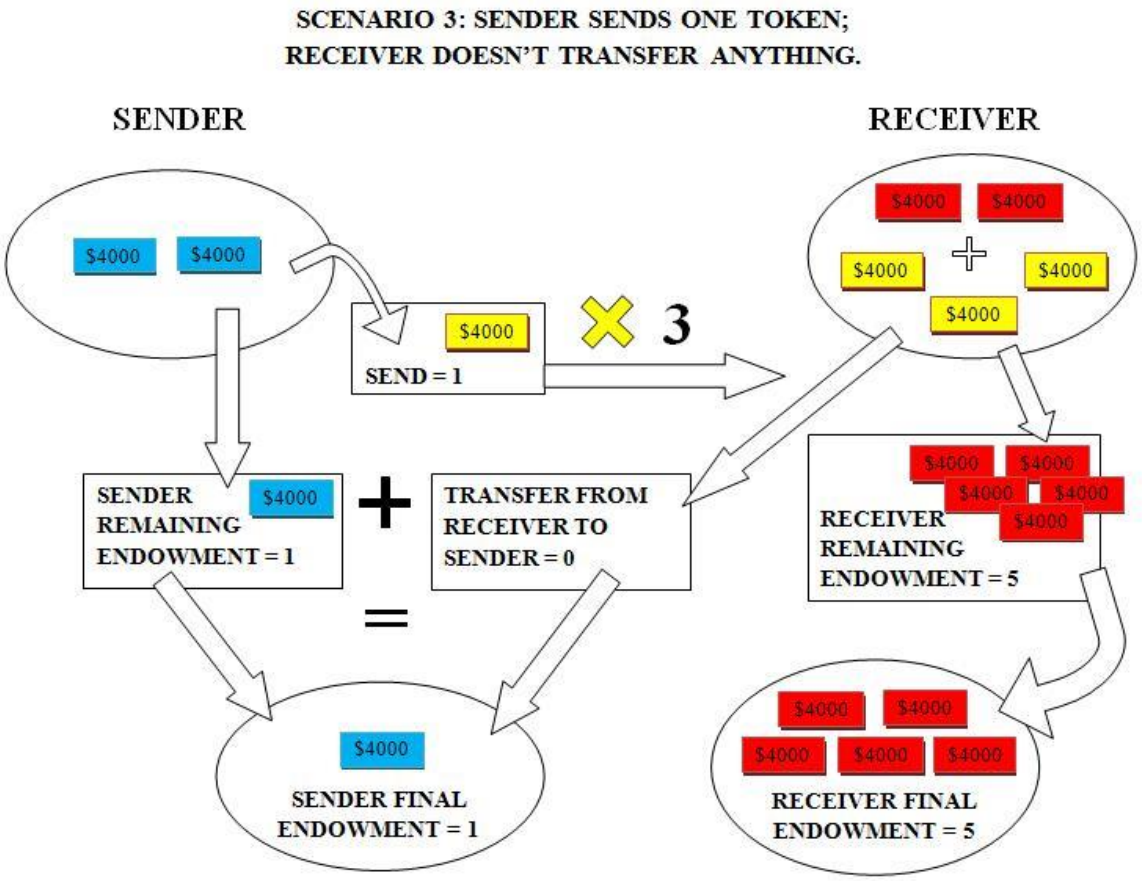
3) *[The instructor explains the interaction on the blackboard while speaking]* This is the representation of the interaction. At the beginning, the Sender and the Receiver receive two tokens. Each token is worth $4000 Colombian Pesos (COP).
4) There are two stages in this interaction. In the first stage, the Sender will make the decision; in the second stage, the Receiver decides. The final payoff is determined according to the decision of both players assigned to the same group.
5) First, we will explain the decision of the sender. The Sender decides whether to send 0 tokens, 1 token or 2 tokens to the Receiver.
6) If the Sender doesn't send anything, the interaction is over and both (Sender and Receiver) remain with the money assigned at the beginning of the interaction. In other words, two tokens for the Sender and two tokens for the Receiver.
7) If the Sender sends one token to the Receiver, the researchers will add two more tokens, so that the receiver is going to get three additional tokens.
8) If the Sender sends two tokens to the Receiver, the researchers will add four more tokens, so that the receiver will get six additional tokens.
9) In other words, the Receiver will always get three times the number of tokens sent by the Sender.
10) Then it is time for the Receiver to make her decision. The Receiver has to decide whether to transfer or not some of the tokens. If she does not transfer, the Receiver keeps everything she has at the end of first stage. If the Receiver decides to transfer, then Receiver and Sender will end up with exactly the same amount of tokens.
11) Let's explain this again. If the sender sends **1 token**, the sender owns one token and the Receiver owns **5 tokens** at the end of the first stage. If the Receiver doesn't transfer anything, they both will end up with this allocation. The Sender's final endowment is one token, while the receiver's one is 5 tokens.
12) If the receiver transfers money, then two tokens will be passed to the Sender. As a result, the Sender's final endowment becomes three tokens (one that he owns plus two that are transferred), and the Receiver's final endowment become of three tokens (five tokens minus two that are transferred).

13) If the Sender sends two tokens, the Sender's endowment at the end of the first stage is zero tokens, and the Receiver's endowment is eight tokens. If the Receiver doesn't transfer, then the final allocations remain the same. The final payoff is **Zero Tokens** for the Sender and the **8 Tokens** for the Receiver.
14) If the Receiver decides to transfer, then four tokens will be passed to the Sender. As a result, the Sender has now a final payoff of **four tokens** (he owned nothing and got four transferred); the Receiver has a final endowment of **four tokens** (eight minus the four passed to the Sender).

Is it clear what the sender and the receiver can do? There are five possible scenarios between Senders and Receivers. These are represented in these graphs.

The assistants distribute the handouts with the graphs of the different scenarios and stress that they can keep them. The instructor illustrates again the five possible scenarios of the interaction, referring to the graphs.

Figure 6: An example of the graphs with the scenarios of the interaction.



15) Now we are going to see how the sender decision-sheet looks like. Please open the envelope and take off sheet number three. Please don't write anything because you still don't know if you will be Sender or Receiver. The Sender has to indicate if she/he wants to send zero, one or two tokens.

SHEET 3

Code:

Please indicate how many tokens you want to send to the receiver. Remember that each token is equal to $4000 COP. Mark clearly just one of the options of the table below.

0	1	2

16) Now let's look at the receiver's decision-sheet. Please take off sheet number 4. Please don't write anything because you still don't know if you will be sender or receiver. The Receiver has to indicate her decision in the next table. The Receiver doesn't know how much the Sender decided to send her, so she has to make a choice for each of the two possible situations. Please mark just one of the two options in each row.

SHEET 4

Code:

For each scenario, mark with a cross if you want or not to transfer some tokens to the receiver:

Scenario 1: If sender sends meone token, then:

I don't transfer anything, we will end up with: - **1 token** ($4000 COP) for the sender; **- 5 tokens** ($20000 COP) for me ☐	I transfer some tokens. We will end up with: **- 3 tokens** ($12000 COP)for the sender; **- 3 tokens** ($12000COP)for me ☐

Scenario 2: If sender sends me two tokens, then:

I don't transfer anything, we will end up with: - **0 token** ($0 COP) for the sender; **- 8 tokens** ($32000 COP) for me ☐	I transfer some tokens. We will end up with: **- 4 tokens** ($16000 COP)for the sender; **- 4 tokens** ($16000COP)for me ☐

Examples

Now let's look at some examples. Please try to answer these questions individually. Then we are going to explain the correct answers on the blackboard. Your answers do not affect the payments or the assignment to the role of the sender or the receiver.

The instructor asks participants to take off sheet number one, where the example questions are given. The instructor leaves about 4 minutes, then the assistants collect the answers and the instructor shows the solutions on the blackboard.

SHEET 1

Code:

Q1. The sender sends one token to the receiver; the researcher adds two tokens. The receiver transfers two tokens to the sender.

1) What is the final endowment of the receiver? _____

2) What is the final endowment of the sender? _____

Q2. The sender sends one token to the receiver; the researcher adds two tokens. The receiver doesn't transfer tokens to the sender.

1) What is the final endowment of the receiver? _____

2) What is the final endowment of the sender? _____

Q3. The sender sends two tokens to the receiver; the researcher adds four tokens. The receiver transfers four tokens to the sender.

1) What is the final endowment of the receiver? _____

2) What is the final endowment of the sender? _____

[The instructor asks them to take off sheet number two].

In this case too, these questions are only used to check that you have understood. These questions do not affect payments or your assignment to the role of sender or receiver. Please, if you don't understand something, raise your hand.

SHEET 2

Code:

Q1. Suppose that the sender doesn't send any token. What will be the sender and receiver final endowment?

A) Two tokens.

B) Zero tokens.

C) One token.

D) It depends on the decision of the receiver.

Q2. Suppose that the sender sends two tokens and the receiver doesn't transfer anything. What will be the sender's final endowment?

A) Two tokens.

B) Zero tokens.

C) One token.

D) Eight tokens.

Q3. Suppose that the sender sends one or two tokens to the receiver. What will be the final endowment of the sender?

A) Certainly more than two tokens.

B) Certainly less than two tokens.

C) It depends on the receiver's decision.

Q4. Suppose that the sender sends one or two tokens to the receiver. What will be the final endowment of the sender?

A) Certainly more than two tokens.

B) Certainly less than two tokens.

C) It depends on the receiver's decision.

Participants are asked to answer the questions in Sheet 2 individually, and to raise their hands when finished. An assistant goes to their place and records the number of correct answers. If all answers are correct, the sheet is collected. If some errors are present, the assistant asks the participant to answer again and to raise again her hand when finished. If the answers are now correct, the sheet is collected. If some errors remain, the lead experimenter illustrated the interaction again and asks the participant to answer again. Only participants who answer correctly the comprehension quizzes are admitted to the experiment.

Ok, now we are going to start. First we are going to draw who will be sender and who will be the receiver. In this bag there are two little cubes: a red one and a black one. We are going to draw one. If we draw the red one, the "even codes" will be the senders and the "odd codes" will be the receivers. If we draw the black one, it will be the other way around. Please remember that the matching between pairs is random.

The instructor asks a volunteer to draw a cube. Senders have to use sheet 3 and receivers have to use sheet 4. The instructor leaves them time to decide. Afterwards the instructor asks them to fold the sheet with the decision inside and the decision sheets are collected. The assistants start entering decisions at the computer.

DECISION NUMBER TWO

Now we move to decision number two. It will follow the same rules as before. We have a sender and a receiver. Those who were senders now will be receivers and vice versa. Don't forget that the matches will be different, so your sender counterpart will be different from your receiver counterpart.

Senders will have to use sheet 3 and receivers will have to use sheet 4. The instructor leaves them time to decide. Afterwards the instructor asks them to fold the sheet with the decision side being non-visible to the researchers, and the decision sheets are collected. The assistants start typing decisions at the computer.

Prior to the session it had been determined randomly if the session is run under the baseline condition or the voting condition. In the former case, the instructor skips the following

section and goes directly to the third decision; in the latter case, it reads the following instructions.

[Voting condition only]:

Please take sheet number five.

Now we are going to determine what you think is the most appropriate decision to take. In this sheet you will find answers for both senders and receivers' decisions. You have to mark your answer for each situation. Later we will state the decisions considered appropriate according to the majority of you. This decision will not determine any payment.

SHEET 5

Code:

Please choose how many tokens you think the sender should send to the receiver. Remember that each token is equal to $4000 COP. Please mark just one of the options bellow.

0	1	2

For each scenario below, mark if it is appropriate or not for the receiver to transfer some tokens:

Scenario 1: If the sender sends one token, then:

The receiver does not have to transfer anything, they will end up with: - **1 token** ($4000 COP) for the sender; **- 5 tokens** ($20000 COP) for the receiver; ☐	The receiver has to transfer some tokens, they will end up with: **- 3 tokens** ($12000 COP)for sender; **- 3 tokens** ($12000COP)for me ☐

Scenario 2: If the sender sends two tokens, then:

The receiver does not have to transfer anything, they will end up with: - **0 token** ($0 COP) for the sender; **- 8 tokens** ($32000 COP) for the receiver; ☐	The receiver has to transfer some tokens, they will end up with: **- 4 tokens** ($16000 COP)for sender; **- 4 tokens** ($16000COP)for me ☐

When they have finished, the instructor asks them to fold the sheet, taking care that decision side is not visible to researchers. Sheets are then collected. Assistants type the voting data into the computer to calculate automatically the results.

Ok, according to your decisions, the majority of you considers appropriate for the sender to send *[report outcome of decisions].*

According to your decisions, the majority of you considers appropriate that the receiver TRANSFER/DOESN'T TRANSFER when sender sends one token.

According to your decisions, the majority of you considers appropriate that the receiver TRANSFER/DOESN'T TRANSFER when sender sends two tokens.

All conditions:

DECISION NUMBER THREE

First we are going to draw who will be sender and who will be the receiver. In this bag there are two little cubes: a red one and a black one. We are going to draw one. If we draw the red one, the "even codes" will be the senders and the "odd codes" will be the receivers. If we draw the black one, it will be the other way around. Please remember that the matching is random. Please remember counterparts are not the same as in the previous two decisions.

[Voting condition only]:

Please remember that you have to choose individually and the previous voting is not binding.

[All conditions]:

The instructor asks for a volunteer to draw a cube. Senders have to use sheet 6 and receivers have to use sheet 7. The instructor leaves them time to decide. Afterwards the instructor asks them to fold the sheet and the decision sheets are collected. The assistants start typing decisions at the computer.

DECISION NUMBER FOUR

Now we move to decision number four. It will follow the same rules as before. We have a sender and a receiver. Those who were senders now will be receivers and vice versa. Don't

forget that you will matched to a different person than your previous matches, so your sender counterpart will be different from your receiver counterpart. Please remember counterparts will be different from the ones in previous decisions. Please remember that you have to choose individually *[Voting condition only]:* and the previous voting is not binding.

[All conditions]:

Senders have to use sheet 6 and receivers have to use sheet 7. The instructor leaves them time to decide. Afterwards the instructor asks them to fold the sheet and the decision sheets are collected. The assistants start typing decisions at the computer.

Now we are going to draw which of the 4 activities is going to be paid. In this bag we are going to put four little cubes with numbers from one to four and we need a volunteer to draw the cube.

The instructor asks them to take sheet number 8. While they fill in the questionnaire, payments are calculated, money is put into envelopes. The assistants write the ID-number on the front of the envelope. A receipt is inside the envelope. Participants must sign the receipt without the researchers being present, and put it in a large envelope stuck on the exit door as they leave the room.

SHEET 6

Code:

Please indicate how many tokens you want to send to the receiver. Remember that each token is equal to $4000 COP. Mark clearly just one of the options of the table below.

0	1	2

SHEET 7

Code:

For each scenario, mark with a cross if you want or not to transfer some tokens to the receiver:

Scenario 1: If sender sends me one token, then:

I don't transfer anything, we will end up with: - **1 token** ($4000 COP) for the sender; **- 5 tokens** ($20000 COP) for me ☐	I transfer some tokens. We will end up with: **- 3 tokens** ($12000 COP)for the sender; **- 3 tokens** ($12000COP)for me ☐

Scenario 2: If sender sends me two tokens, then:

I don't transfer anything, we will end up with: - **0 token** ($0 COP) for the sender; **- 8 tokens** ($32000 COP) for me ☐	I transfer some tokens. We will end up with: **- 4 tokens** ($16000 COP)for the sender; **- 4 tokens** ($16000COP)for me ☐

Questionnaire (extract)

Q1) Sex:

M	F

Q2) Age: _____

Q3) Marital status:

Married/Free union	Single	Divorced/Widower

Q4) Normally, are you fully prepared to take risks or do you try to avoid them? Please answer using the scale below from one to five, where one means "fully prepared to take risks" and five means "fully unfit to take risks".

Fully prepared to take risks	1	2	3	4	5	Fully unfit to take risks

Q5) Normally, are you fully prepared to take financial risks or do you try to avoid them? Please answer using the scale below from one to five, where one means "fully prepared to take risks" and five means "fully unfit to take risks".

Fully prepared to take risks	1	2	3	4	5	Fully unfit to take risks

Q6) Including you, how many people live at your home? _____

Q7) According to the utility bills, which is your socioeconomic status?

1	2	3	4	5	6

Q8) What is your education level?

Elementary ☐

High school ☐

Some university, but not graduated yet ☐

Technical ☐

University degree ☐

Q9) What is your occupation?

Director or manager ☐

Professional scientist or intellectual ☐

Technician or medium level professional ☐

Administrative support personnel ☐

Utilities worker or salesman ☐

Agriculturist or agrarian, forestall and fishing qualified worker ☐

Officer, operative or mechanical arts craftsman ☐

Facilities and assembly machines operative ☐

Elementary occupation ☐

Military occupation ☐

Unemployed, retired or housekeeper ☐

Student ☐