Endogenous Institutions: a network experiment in Nepal

Giulio Iacobelli[*] Juni Singh [Job Market paper][†]

October 8, 2020

Abstract

In developing countries where formal institutions are often weak, the community is responsible to enforce local agreements. In such settings, peer monitoring represents a natural mechanism for the enforcement of agreements. This paper studies the demand for monitoring and its effectiveness in sustaining cooperation across social groups. We map social networks of 19 villages in rural Nepal and conduct an experiment to explore the role of the endogenous choice of monitors. Individuals play in groups of three, both with their close friends and with people socially distant in their network. They receive the opportunity to anonymously choose their preferred "institution". We combine a theoretical model and a unique lab-in-the-field experiment to show that closely knit groups are significantly more likely not to choose a monitor, while sparse groups tend to prefer a monitor who is highly central in the network. The endogenous selection of monitoring ensures higher cooperation compared to an exogenous assignment, but only in sparse groups. Further, we observe that the outcome of the vote acts as a signal of intra-group trust.

Key words: Network, Peer Monitoring, Experiment, Public good game, Sparseness
JEL Classification: C93, L14, P48

We would like to thank CEPREMAP, LABEX OSE and Sarah Andrieux Fund for the funds to conduct the field experiments. We thank our partner organization, Rooster Logic Nepal and the enumerators for the execution of the field project. We are grateful to Philippe Jehiel, Abhijit Banerjee, Margherita Comola , Nicolas Jacquemet and Philipp Ketz for their tremendous guidance. We have benefited from the comments of Francis Bloch, Yann Bramoulle, Beatrice Boulu Reshef, Emily Breza, Vessela Daskalova, Fabrice Etile, Edoardo Gallo, Gabrielle Demange, Esther Duflo, Ben Golub, Cynthia Kinnan, Sylvie Lambert, François Libois, Jean François Laslier, Karen Macours, Emanuela Migliaccio, Paolo Santini, Paul Seabright, Lennart Stern and Karine Van Der Straeten. In additon, we acknowledge financial support from the Belgian French speaking community ARC project n°15/20-072 of Saint-Louis University - Brussels.

[*] PhD candidate at Paris School of Economics

[†] PhD candidate at Paris School of Economics

1 Introduction

The engagement of local communities is a standard mechanism for channeling development programs. The benefit of this approach lies in increased participation of communities and in deeper adaptation to local needs. Yet, the efficiency of such initiatives crucially depend on the institutional environment and on the level of cooperation. Development practitioners recognize the increased risk of free riding when institutions are imposed without involving the community (Mansuri and Rao, 2004; Narayan, 1995). In this context, it is important to understand which informal institutions sustain and promote cooperation. Peer monitoring represents a natural mechanism for the enforcement of social norms in such settings. Previous research has focused on the effectiveness of peer monitors but the demand for such a mechanism is understudied. Our paper fills this gap in two main directions. First, it studies the demand for peer monitors across social groups. Second, it explores how endogenous monitors affect cooperation.[1] We focus on the role of social networks in the choice of these institutions and its impact on cooperation.

In developing countries, given the weakness of formal institutions, social networks represent a relevant tool to regulate transactions between individuals. Communities mostly rely on interactions with socially proximate peers and third party institutions are often used to enforce contracts among socially distant individuals. Ostrom, 1991 suggests how mechanisms based on social ties have sustained cooperation[2] rather than mere punishment.[3] We conduct a lab in the field experiment in 19 villages in rural Nepal to understand third party monitoring and cooperation. Villagers play a cooperation game and can choose to elect a monitor through majority voting. This monitor can impose higher cooperation through reputational concerns. Concerns of social image are important drivers of cooperative behavior since people fear bad reputation. Third party monitoring can substitute for social density by emphasizing these very reputational concerns (Greif, 1989; Bowles, 2008; Andreoni et al., 2020). Further, the impact of the monitor may depend on his position in the social network[4] and on the political process whereby it is assigned to groups. We explore the

---

[1] See Olken, 2007; Bjorkman et al., 2009 and Gelade, 2018. Unlike direct reporting in these papers, our experiment relies on reputational concerns affected via gossip by the chosen monitor.

[2] In rural Nepal, community-based organizations (Forest User Associations, Water User Associations and Cooperatives) present a good example of the success of such reputational mechanisms.

[3] There is a huge literature that shows the effect of punishment on public good games. For more details see Fehr and Gachter, 2000; Charness et al., 2008; Beersma and Kleef, 2011; Wu et al., 2015, 2016; Sommerfeld et al., 2007; Glockner et al., 2007; Galbiati and Vertova, 2008; Kosfeld and Rustagi, 2015; Fonseca, 2018 and Fehr and Sutter, 2019.

[4] Central individuals in the network are shown to be particularly effective in monitoring due to their higher ability of spreading information in the form of gossip (Ballester et al., 2006 and Banerjee et al., 2019)

effects of monitoring induced by reputation concerns[5] rather than by material punishment. Two main pieces of evidence emerge from previous research. First, dense groups are able to sustain more cooperation than socially distant groups. Second, monitors have the power to relax the inefficiencies arising from contractual incompleteness in the context of socially distant groups (Breza and Chandrashekhar, 2019; Breza et al., 2016; Chandrashekhar et al., 2018). These studies focus on the impact of exogenously assigned monitors on cooperation in groups.

The aim of this paper is to bring the literature forward by allowing individuals to endogenously elect their preferred monitor and by studying the induced cooperative behavior. We estimate the demand for monitoring, relate it to the network structure and study its impact on cooperation by tackling three subsequent questions. First, do individuals change their demand for monitoring as a function of the social composition of the group they interact with? Second, do monitors who are endogenously chosen spur cooperative behavior compared to those assigned exogenously? Third, is the election of monitors perceived as a signal of intra-group trust? To answer these questions, we conduct a lab in the field experiment in rural Nepal and build a theoretical model supporting our experimental findings.

First, we ask whether groups with different social proximity elect different third party institutions. In line with the literature, we offer three monitoring options, according to a measure of social prominence. We present strong evidence that socially distant groups are more likely than closely-knit groups to elect a high central monitor. We find that individuals are 40% less likely to elect a monitor with their close peers compared to when they are in groups with socially distant members. This supports the idea that contractual incompleteness can be mitigated by social density, but socially distant individuals need third party institutions to enforce social norms and increase efficiency.[6]

Secondly, we investigate whether the political process by which the institution is chosen matters for cooperative behavior. Interestingly, we find that a monitoring institution that is democratically elected has strong positive effects on cooperation compared to an institution assigned randomly. Previous experimental evidence in economics (Sutter et al., 2010; Tyran and Feld, 2006; Dal Bo et al., 2010) and sociology (Grossman and Baldassarri, 2012)

---

[5] Individuals rely on local connections for risk sharing, public good provision and information delivery. Social image and trust therefore become more important. See Kranton 1996; Leider et al., 2009

[6] In line with this reasoning Glaeser et al., 2000 establish that groups with shorter social distance have higher trust. In the same spirit, experiments in the lab (Hoffman et al., 1996; Leider et al., 2009; Goeree et al., 2010) and in the field (Etang et al., 2011; Baldassarri and Grossman, 2013) show that cooperation increases with decreasing social distance.

show that cooperation is higher when players are given the opportunity to choose the institution rather than having an externally imposed one. We dig deeper in this dimension, and we offer evidence that the positive impact of endogenous institutions is limited to socially distant groups. More precisely, the magnitude of the increase in contribution ranges from 8.6 % when a socially prominent monitor is chosen to 21.9 % when no monitor is chosen. In socially close groups, the point estimate is negative and not significant. The possible mechanism underlying this "democracy premium" can be explained by an increased sense of agency and control, increased sense of authority (Greif, 2006) and stronger worthiness of authority (Zelditch, 2001).

Third, we present the first evidence of heterogeneous impact of asymmetric information in different social groups. The theoretical literature of principal-agent models (Herold, 2010) shows that the proposal of a complete contract can signal distrust. The principal may thus prefer to leave the contract incomplete rather than to signal distrust by proposing a complete contract. On this basis, we explore if group members perceive the election outcome as a signal of intra-group trust. Individuals in closely-knit groups are particularly affected by the outcome of the vote. We find that the group interprets the outcome of the vote as a strong signal of trust when no monitor is chosen. This entails a significant increase in contribution by 19.7% in sparse groups. On the contrary, dense groups perceive the election of a high central monitor as a signal of mistrust leading to a decrease in contribution by 13.3%. This finding sheds light on how the effects of signaling are heterogeneous when we take into account the social structure of the agents.

To answer these questions, we first conduct an intensive network survey inspired by Banerjee et al., 2013. We ask questions about advice, trust, friendship and financial relationships such as "who do you spend your free time with", "in case of an emergency, who would you rely on" and "who do you borrow money from". Based on data from the survey, we build an undirected social network where a connection between two people is established if any one of them names the other. Next, we identify the person in the network with the highest and lowest eigenvector centrality. In other words, we identify the most influential and least influential individuals in the village. The rest of the village is divided into groups of three with varying network distance. Groups of three are an optimal choice to study questions of mutual trust inasmuch as it is possible to maximize the behavioral contrast between groups which can sustain high levels of cooperation and those which cannot (Jackson et al., 2012). Players interact both in a closely knit group and in a group with socially distant individuals. The closely knit groups are often homogeneous in terms of caste.

We refer to the closely knit group as “dense” and to the group with acquaintances as “sparse”. In formal network the terms “dense” implies each individual is at most at distance 2 (average path length < 1.6) and “sparse” implies each individual is at least at distance 4 (average path length > 4). We allow individuals to either vote for no monitor, or to choose one from two monitor candidates belonging to their village: a high central monitor (very prominent individual) or a low central monitor (less prominent). Players play a contribution game both with a monitor chosen by the group (endogenous treatment) and an externally assigned monitor (exogenous treatment). The monitor does not materially punish but only observes the contributions of each player, which would otherwise be private information. Players have an initial endowment and need to decide between how much to contribute to a common pot and how much to keep for themselves. The total contribution in the common pot is augmented by 50 percent and divided equally among the three players irrespective of initial contribution. Each individual therefore plays both in a dense and in a sparse group, and in each group both with exogenous and endogenous monitoring institutions. The order of all treatments is randomized.

We vary both the social composition of groups and how monitors are assigned to study the demand for monitoring and cooperation. The rest of this paper is organized as follows: Section 2 presents the experimental protocol and the data collection process. Section 3 describes the results of the experiment and the econometric specifications we use. In Section 4, we build a theoretical model supporting our empirical results. We discuss the results and conclude in Section 5.

2 Experiment

2.1 Overview: Networks and Data

We start by mapping the social network of villages, with a special focus on relations of trust. Given the location of these villages[7], mutual trust fundamentally shapes social interactions and the contribution to local public goods. As a first step, we assigned a unique identification code to each woman in the census. We started interviewing very few individuals, who would give us names of their closest friends and we administered the network questionnaire to those women who were nominated in the first round. This process was repeated iteratively until either all women were covered or no new individual was nominated - the elicited network is "closed". This technique has the advantage to be faster than the standard net-

---

[7] The villages are situated at 1200m above sea level in the mid hills of Nepal. They are a four hours drive away from Kathmandu

work elicitation method and simplifies considerably the issue of homonyms. Each woman was asked at least three connections for each question. The questionnaire consisted of a set of questions designed to elicit social networks, inspired by Banerjee et al., 2013. These questions are meant to elicit ties of friendship and trust and span along various dimensions of social interactions. A link between two individuals i and j is established when either i nominates j or vice versa in any of the questions. We then aggregate and collapse the networks obtained from different questions into one, undirected network. Once a network is fully mapped, it is possible to visualize it and extract important statistics that are central in our experimental design. Figure 1 in the Appendix is a snapshot of the network of a village where we conducted the experiment.

The network we obtain is thus a good representation of the social structure of the community and it is an essential variable of our study. More precisely, we use the network to create groups of contrasting social density for every participant and interact it with variations along two dimensions: monitoring centrality and the political process by which monitors are assigned, either democratically elected or exogenously given. We focus on networks of only women due to the high emigration rate of men either to Kathmandu or abroad, as shown by our pilot experiment conducted in the spring 2018. In the districts we worked in, social networks are often gender specific and women play a preponderant role: they are responsible for households' finances, for agricultural production and for their children.

We look at how social density influences the demand for monitors and how it ultimately affects individual contribution to public goods. Groups are formed in order to maximize the number of participants who play in both dense groups, i.e. groups of average path length less than 1.6, [8]and in sparse groups, i.e. with average path length higher than 4. In other words, being in a dense group implies that the members of the group are at no more than 2 steps away from each other whereas in the sparse group they are at least 4 steps away. The cutoffs defining dense and sparse have been carefully chosen in order to amplify the respective contrast in trust and reputation while maximizing the number of observations. The starker is the difference between dense and sparse groups, the more different will be the behavioral response in the different treatments. Figure 7 in Appendix B shows the distribution of average path length of all groups we formed. We over sampled dense groups to make a reliable comparison with sparse ones. Players in the dense groups often belong to the same caste and have similar characteristics. We end up with 503 women who played

---

[8] The dense groups would correspond to topography that is a triangle (average path=1), line (average path=1.3)

in both sparse and dense groups, as defined by our thresholds. The summary statistics are presented in Table 7. In total, we have four observations for each participant, for a total of 2012 observations.

We choose monitors candidate with respect to their Bonacich centrality and their assignment to groups can be determined by either democratic election or random exogenous assignment. In order to neatly disentangle the different possible channels that might drive behavior, we set up an experiment where groups of three individuals are asked to privately vote for their preferred monitor and then play twice a standard public good game. The experimental session is sequenced as follows: first, players are assigned to a group formed either by their closest friends or by socially distant peers. The order of assignment to these two group compositions is randomized. Secondly, after being assigned their groups, players privately vote for their preferred monitor. Third, the choice of monitor is immediately followed by a contribution game. Each individual plays 2 rounds of a public good game within each group, once played with the elected monitor and once with a randomly picked monitor option, where we randomize the order of the two treatments. Groups are then reshuffled so that the same player is then placed in a different group composition (dense or sparse) and the game unfolds again as explained above. In total, each individual plays 4 rounds in two different groups (dense and sparse).

After participants play in the experimental sessions and receive payment for their performance in the games, we administer a second questionnaire meant to capture caste, wealth, religion, membership to community based organizations and a set of other individual level characteristics. Participants are quite homogeneous in terms of wealth and networks display high level of clustering with respect to caste.

2.2 Experimental context

Nepali villages are often too remote to be reached easily or too sparse to ask their members to participate to group sessions in a fixed location. We decided to conduct our experiments in the mid-hills of Nepal in the district of Makwanpur, which is around four hours drive from Kathmandu. The municipalities we chose - Palung, Bajrabarahi, and Chitlang - present an economy almost uniquely focused on agriculture and the exploitation of natural resources. Dozens of community based organizations are active in the region and people are generally involved in at least one, they are familiar with issues of coordination and with the risks of free-riding. Villages are on average composed of 70 households, for an average of 120 women per village. We covered 19 villages with more than 2000 women between

18-60 years answering our network survey. We have a census of all inhabitants living in each village and we made sure to administer the network questionnaire to every woman.

In partnership with a local research company based in Kathmandu, we hired a team of local enumerators. All enumerators were women, in order not to add any confounding factor in the network elicitation and in the experimental sessions. In each village, women who answered our network survey were invited via a phone call to take part in the experiment. We invited around 75% of interviewed women in each village and, based on a measure of (eigenvector) network centrality, we divided people into either players or monitor candidates. The individuals belonging to the top and bottom 5% of the centrality distribution were assigned the role of monitor candidates, while the others were assigned the role of players. Among those who were assigned the role of players, we oversampled groups in the periphery so as to avoid picking high centrality individuals in order to maximize the contrast between dense and sparse groups. This gives more power to the information transmission role of the monitor. As an incentive to participate, every player was given 100 Nrs (1€) along with the possibility to obtain additional money up to 200 Rs, as a function of their performance in the games. On average, the total gain was around 220 Rs. per individual which is half day's wage. Monitors were given a fixed sum of 250 Rs. for their participation.

The experiments were typically conducted early in the morning in schools close to each village. Women, as they arrived to schools, were assigned to either sparse or dense groups for the experiment. They were progressively sent to one of the classrooms to play the games. Once played, they got out of the room to be assigned to another group and to play again with a different group composition. The order of the dense and sparse groups was randomized. Typically, three sessions were run in parallel in separate classrooms with one session lasting for around 15 min. Two enumerators were in charge of each session: they read the instructions, conducted the game and noted down the choices of participants.

2.3 Design

In our experiment, we have three treatments variables. First, group composition. Groups can be composed either by close friends or by people socially distant in the network. Second, centrality of monitors. In our experiment, we offer three monitoring options: high central monitors, low central monitors and no monitors. Third, the process whereby monitoring institutions are assigned: either democratically elected by the group or exogenously imposed. After assigning the role of high central and low central monitors, which remains fixed throughout the experiment, we divide the rest of the individuals into groups of three with varying group composition, either dense or sparse. Individuals play in groups of three in both dense and sparse treatment in a randomized order. In Figure 1, we show two possible groups for the player circled in green. She plays both with her closest friends - circled in red - and with individuals far in the social network circled in blue. By always reshuffling groups in such a way that every individual plays exactly in two different groups, we are able to extract individual fixed effects. This part of the design is of paramount importance because of the intrinsically endogenous nature of networks: the network position of player i is endogenous to her observable characteristics which are in turn affecting her contribution. This design allows a neat disentanglement of the endogenous position in the network from the contribution, through the extraction of fixed effects at the individual level.

At the start of each session, group players are gathered in a room where they can see each other, but no communication is allowed. Each member of the group receives 10 tokens of a different color, where the value of 1 token is marked at Rs 10. Each session is divided in two stages. In the first stage, each player privately casts a vote on her preferred monitoring option.[9] In Figure 2, step 1 represents the setting of the game. Players are given the option to choose between high central monitor (H), a low central monitor (L) or no monitor at all (NM). Note that this monitor is a fourth "player" that remains the same for all groups within a village. The cost of choosing the monitor is 20 Rs.[10] This cost makes always choosing a monitor a non - dominated strategy. The cost is paid by participants who vote to have a monitor (either high central or low central), irrespective of the voting outcome of the group.[11] The monitor is elected by a majority rule and the result of the vote is not immediately revealed. As seen in Step 2 of Figure 2, the group is then randomly assigned to either the endogenous treatment or the exogenous one. The randomization is implemented by picking one out of two balls: if the ball drawn is green, the endogenous treatment is played first and the exogenous follows. If the ball drawn is pink then exogenous is played first followed by endogenous. The result of voting is only revealed just before playing the endogenous treatment. In the exogenous treatment, the group is randomly assigned either to a high central, low central or to no monitor treatment.

In the second stage of the experiment, the group plays a public good game where each player decides how many tokens out of the 10 is to be contributed to the public pot. They

---

[9] In case of a tie, the monitor choice was determined by a random draw. Ties represent around 6% of cases.

[10] In line with the public good literature, the cost of the option was around 7% of the average earnings across all games.

[11] If x votes for a monitor but no monitor is elected by the group, x stills pays the cost of voting for a monitor.

are informed that the money in the public pot would be increased by 50% and then divided equally among them. As seen from Step 3 of Figure 2, once the contributions are made, the monitor -either elected or assigned- is called into the room to see how much each player contributed in the public pot. The monitor can distinguish the contributions belonging to each player by the different colors of the tokens they were endowed with. Moreover, the monitor does not have the power to impose fines and simply observes how much each player contributed. We exploit only the informational channel whereby the players' reputation can be affected (e.g. gossips, reporting etc.), following the assumption that it would drive a lot of real-life interaction in the village. We study how the fear of being reported on by the monitor outside the lab drives the behavior of people and how it consequently affects the demand for third party monitoring[12]. To sum up, the contribution game is played twice in the same group without receiving any feedback, once with the monitor option chosen by the group (endogenous) and once randomly assigned monitoring option (exogenous).

3 Results

The hypothesis is that the individual demand for peer monitoring varies depending on the identity of group members. In particular, we expect individuals in dense groups to not choose a monitor and to enforce co-operation on their own in the second stage of the game. This result would not hold for socially sparse groups where the temptation to free ride is higher, given the lower level of reciprocal trust. Thus, socially sparse groups might have a stronger incentive to pay the fixed cost of electing a monitor that is able to strengthen the reputation channel and the level of intra-group trust. The presence of a monitor - even more so for a high central one - increases the possibility of being reported on outside the lab in case of "defection". On the other hand, we expect socially close groups to be co-operative irrespective of the treatments[13]. We also expect to find a different impact between the endogenous and the exogenous assignment monitors. The choice of the group is revealed only in the endogenous treatment, since it can affect contributions by carrying additional information about intra-group trust.

---

[12] Breza et al., 2016 do not find a significant difference between information and punishment treatments

[13] Socially close or distant is characterized by the average social distance (path length) in a group

3.1 Preliminary findings and limitations

We start the analysis by looking at the individual level variation in the choice of the monitor. In Table 1, the numbers along the diagonal represent the percentage of individuals that always choose the same voting strategy irrespective of group composition. The largest proportion being 34.95% that always chooses to have no monitor, followed by 19.68% that always vote to have a high central monitor. The voting result shows substantial variation in voting strategy. Looking at the aggregate demand for peer monitoring, both dense and sparse groups vote more often to not have a monitor. Figure 3 shows that in dense groups, around 32% of players vote for a high central monitor, while in sparse groups more than 39% of players do so. Low central monitor is seldom chosen accounting for around 13% in both dense and sparse groups. For contribution, exogenous monitoring increases contribution only in sparse groups as seen from Table 2. We want to study how this differs when individuals have the power to demand their own monitoring institution. To begin with, we compare the outcomes under endogenous and exogenous institutions, clubbing all three monitor treatments together for the later in Figure 4. ‘How’ the monitoring institution is obtained matters only for sparse groups where endogenous monitoring in blue increases contribution compared to the exogenous one.

Before presenting the results, we want to highlight a possible limitation of the result given the caveats in the process of network elicitation. We ask for at least three "nominations" of friends. In most interviews, women named an average of 3-4 women which may not be fully exhaustive and may lead to networks that are much sparser than they actually are. This could imply an overestimation of social distance, i.e. individuals are actually socially closer than what they appear to be, which in turn may bias the result. It does not however represent a threat to the validity of our results. On the contrary, it implies that the estimated effects of our treatments represent a lower bound of the real effect.

3.2 Statistical Estimation

3.2.1 *Impact of group on monitor voting: Sparse groups elect a high central monitor more often than dense*

As suggested by the preliminary results shown in Figure 3, we conduct a Mann-Whitney test to understand whether the proportion of participants choosing a given monitor is significantly different across group compositions. We find that no monitor is chosen significantly more often in dense groups rather than in sparse groups (p-value 0.07) and that high central monitors are chosen more often in sparse rather than in dense groups (p-value 0.002).

In order to estimate how the demand for monitor varies depending on the group composition, we use a multi logit regression with individual and round fixed effects. Since players vote once in a dense and once in a sparse group in a random order, we can include both individual and round fixed effects, therefore exploiting a "within" design. The fixed effect multi logit model is therefore defined by the logistic probability of choice of monitor y_{jt}, where y_{jt}=0: No monitor, y_{jt}=1: Low central monitor and y_{jt}=2: High central monitor. We take y_{jt}=0 as base category and can write the fixed effect logit as

$$Pr(y_{jt} = 1) = \frac{1}{1 + e^{-(\alpha+\beta_1 G_{jt}+\beta_2 X_g+\rho_j+\nu_t+\epsilon_{jt})}}$$

$$Pr(y_{jt} = 2) = \frac{1}{1 + e^{-(\alpha+\beta_2 G_{jt}+\beta_3 X_g+\rho_j+\nu_t+\epsilon_{jt})}}$$

where y_{jt} is the chosen level of monitoring, G_{jt}: dummy for group composition equal to 1 if the treatment is for dense groups, X_g: group characteristics ν_t: round fixed effect and ρ_j: individual fixed effects.

We present in Table 3 the results of the multinomial fixed effect regression of individual monitor choice (voting) on the social composition of the group (dense / sparse). In the first column, we find that dense groups are less likely to elect a high central monitor by 40% points compared to sparse groups. In the second column, we see that this is also true when we control for group characteristics (if the members have same caste and same education level). More details on voting as a function of individual characteristics is presented in Appendix C. This result is in line with our hypothesis that closer groups would prefer not to have monitoring whereas individuals in sparse groups would want a high level of monitoring.

3.2.2 *Impact of different exogenous monitoring: High central monitor increases cooperation relative to no monitor in sparse groups*

For contribution, we start with the baseline case where monitors are assigned exogenously and study the difference in contribution between sparse and dense groups. As seen from Table 2, in sparse groups, average contributions increase significantly (p-value 0.014) by Rs 7.4[14] (15.8% of the mean) in the presence of a high central monitor (H) as compared to no monitor (NM). In dense groups, there is a Rs 4.5 increase (8.3% of the mean) but the difference is not significant. This result is in line with the literature that suggests presence

---

[14] Note that the value of 1 token is Rs 10. The regression is in terms of tokens but all the results are expressed in terms of Rs.

of a central monitor increases cooperation only in sparse groups (Breza et al, 2016). Further the cost of the monitor being 8% of the average payoff, it is optimal for sparse groups to vote for a monitor but not dense. Taking only the exogenous monitor treatment, we run a linear regression with fixed effects on the contribution with respect to the type of monitor that was assigned and the group composition. It takes the following form:

$$c_{jt} = \alpha + \beta_1 \cdot Dense + \beta_2 \cdot H + \beta_3 \cdot L + \beta_4 \cdot H \times Dense + \beta_5 \cdot L \times Dense + \rho_j + \nu_t + \epsilon_{jt}$$

where c_{jt}: contribution of individual j in round t, $Dense$: dummy equal to 1 if the group is dense, H: dummy equal to 1 if a high central monitor is assigned, L: dummy equal to 1 if a low central monitor is assigned, ρ_j: individual fixed effect and ν_t: round fixed effects.

We are particularly interested in the coefficient β_2 that shows the effect of being assigned a high central monitor and β_4 that shows the difference in the effect across dense and sparse groups. In the first column of Table 4, the dependant variable is the monitor choice. We see that dense groups in general contribute Rs 13.7 higher (23% of the mean) than sparse ones. Next, contribution increases by Rs 7.25 (11% of the mean) in the presence of a High central monitor (H). As seen from the interaction term[15], the effect is starker in sparse groups. The second column shows similar effects when controlled for group characteristics. In both columns we control for the monitor choice of the individual.

3.2.3 *Impact of endogenous v/s exogenous monitoring: Holding the monitoring institution fixed, for a given population, choosing a monitor increases cooperation only in sparse groups*

To answer the impact of "how" the monitoring institution is chosen impacts contribution, we estimate a linear fixed effect regression. In the endogenous treatment, individuals 'select' into an institution that drives their contribution in addition to "how" the monitor is chosen. In order to overcome this selection problem, we keep monitoring fixed and compare groups which play both exogenous and endogenous treatment under the same monitor. Our identification strategy is to overcome selection by comparing the same group, with the same monitor treatment, differing only on how this monitor was obtained. Inspired by Dal bo et al (2010), an individual $i's$ action in the game may depend on the group density G ϵ {dense, sparse}, elected monitor M ϵ {NM, H ,L} , mechanism that selected the monitor I

---

[15] $H \times Dense$ being an interaction term represents [(H=1)-(H=0)|Dense]- [(H=1)-(H=0)|Sparse]

ϵ {Endo, Exo} and her type α_i. The probability of cooperation is therefore determined by

$$P_i = f(M, G, I, \alpha_i)$$

We fix the group G and monitor M to determine the effect of the mechanism by which the monitor is elected. More formally,

$$E(P_i|G = dense, M = NM, \alpha_i, Endo) - E(P_i|G = dense, M = NM, \alpha_i, Exo)$$

By doing so, we eliminate the threat of self-selection and we are able to disentangle the effect of the exogenous vs endogenous treatments. In terms of regression, it translates into the following fixed effect equations.

$$\begin{aligned}
c_{jt} &= \alpha + \beta_1 \cdot (Endo \mid G = S, M = H) + \rho_j + \nu_t + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Endo \mid G = D, M = H) + \rho_j + \nu_t + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Endo \mid G = S, M = NM) + \rho_j + \nu_t + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Endo \mid G = D, M = NM) + \rho_j + \nu_t + \epsilon_{jt}
\end{aligned}$$

where c_{jt}: contribution of individual j in round t, *Endo*: is a dummy variable that takes value 1 if monitor is endogenously chosen , given group $G : \{D = dense, S = sparse\}$ and monitor choice $M : \{NM = No\,monitor, H = High\,central\,monitor\}$,[16] ρ_j: individual fixed effect and ν_t: round fixed effect. We are primarily interested in the coefficient β_1 that captures the effect of having an endogenous monitor as compared to being assigned exogenously.

Figure 5 shows the average contributions for sub samples that are free from selection effect. We see that for a sparse group, contribution increases under an endogenous monitoring setting as seen form the red bars. In particular, with endogenous no monitor, contribution increases significantly (p-value 0.009) by Rs 9.1 while with a high central monitor it increases by Rs 5 but not significantly. The change in dense groups across endogenous and exogenous monitoring institutions is not significantly different. We find in sparse groups that giving individuals opportunity to chose their own monitoring institution leads to better outcomes than externally imposing a third party monitor.

The first column in Table 5 reports results for individuals who self-selected into high

---

[16] We also tried to do individual level analysis by looking at variation in monitor choice within groups. We find that 50% of the groups vote unanimously for the same monitor option hence not much power to study this effect.

monitoring institution in sparse groups followed by dense groups in the second column. The next two columns report the same but for the case where groups self selected into no monitoring. We see that a sparse group electing a high central monitor (H) endogenously increases contribution by Rs 5 (8.6% of the mean) whereas there is no effect for the dense group. Similarly, sparse group electing no monitor (NM) endogenously increases contribution by Rs. 9.13 (21.9% of mean) with no effect in the dense groups. This result presents evidence to believe that there exists a sort of "endogenity premium": individuals facing the same monitoring institution behave differently depending on whether the institution is chosen by the group itself or imposed.

3.2.4 *Impact of order endogenous/exogenous on contributions: Contributions are higher under EndoExo for dense NM and lower for dense H*

Next, we look at the effect of order of endogenous and exogenous treatment on contribution. The result of the vote is only revealed in the Endogenous case, hence if Endogenous is played first, there is a possible extra information[17] that could affect contribution in both the endogenous and exogenous treatment. We focus on No Monitor and High central monitors because of the very few observations we have for Low central monitor. In presenting this comparison, we plot the average contribution in treatments across monitoring condition and group composition that vary by the order in which endogenous and exogenous treatments were implemented. We can see in Figure 6 an evidence of a possible significant effect of the order, especially in dense groups. The election of a high central monitor in dense groups decreases contribution by Rs 9.7 significantly (p-value 0.09). When no monitor is elected, contribution increases by Rs. 7.1 in sparse groups and Rs 9.4 in dense groups (p value 0.07). We run OLS regressions controlling for individual level characteristics. The dependant variable takes value 1 if the endogenous round is played first and 0 otherwise. We look at the average contribution across the endogenous and exogenous treatments conditional on endogenous being played first v/s if it is played second.

$$
\begin{aligned}
c_{jt} &= \alpha + \beta_1 \cdot (Order \mid G = S, M = H) + \beta_i \cdot X + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Order \mid G = D, M = H) + \beta_i \cdot X + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Order \mid G = S, M = NM) + \beta_i \cdot X + \epsilon_{jt} \\
c_{jt} &= \alpha + \beta_1 \cdot (Order \mid G = D, M = NM) + \beta_i \cdot X + \epsilon_{jt}
\end{aligned}
$$

---

[17] We hypothesize that this information could act as a signal of the level of trust in the group vis à vis each other.

where c_{jt}: contribution of individual j in round t, $Order$: is a dummy variable that takes value 1 if endogenous treatment is played first , given group $G : \{D = dense, S = sparse\}$ and monitor choice $M : \{NM = No\,monitor, H = High\,central\,monitor\}$, X: individual characteristics (caste, wealth, age and education). We are primarily interested in the coefficient β_1 that captures the effect of having played an endogenous monitor round first followed by the exogenous one.

In Table 6, we see that the effect of revealing the group's choice has a stronger signalling in dense groups compared to the sparse. Contribution decreases significantly by Rs. 13.68 (13.3 % of mean) in dense groups that played endogenous first and elected a high central monitor (H). On the other hand contribution increases by Rs. 12.73 (19.7 % of mean) for dense groups and Rs. 7.7 (7% of mean) for sparse groups when groups played endogenous first and elected No monitor (NM). We hypothesize signalling is stronger in dense groups, since they have a stronger prior about the altruism level in their group and are more likely to have future interactions. We model this in the section below.

4 The model

Through our data, we observe that the density of groups; either close in the network (dense) or far away (sparse) matters in demand for peer monitoring. Particularly, sparse groups vote more often to have a high central monitor. Through the model, we want to show this difference in demand for peer monitoring. Further, how the monitor is chosen matters for cooperation. The group having the choice to choose it's own monitor increases cooperation when no monitor is elected and decreases cooperation otherwise. The group's vote outcome may act as a signal about the level of altruism and trust in the group. The outcome of the vote being revealed only in the endogenous setting carries extra information that can change the belief of player i about the type of player j. We want to study the possible differences in contribution when endogenous is played first versus played second via a simple two player model where players vary in the level of altruism.

Types

We model the contribution behavior of individuals with an *altruism* parameter α. We can think of this parameter as determining the propensity of higher contribution. As people become more altruistic, i.e, the value of α increases, individuals are more likely to contribute a higher amount. Each player depending on the group can have a level of altruism $\alpha_i \in$

$\{\alpha_l, \alpha_h\}$, where $\alpha_h > \alpha_l$. We further think of the parameter α as determining the choice of the monitor. Player i knows her own level of altruism α_i and forms a prior $\mu_{0i}(\alpha_j)$ on the level of altruism of the other player j. Let us assume that the subjective probability of individuals being type α_h is p in dense groups and q in sparse, where $p>q$[18]. This mimics the fact that in dense groups, people perceive their neighbors to be more altruistic compared to in a sparse group. Let d_i represent group composition where $d_i \epsilon \{dense, sparse\}$.

The initial prior of Player i about j is then as follows:

$$\mu_{0i}(\alpha_j = \alpha_h) = \begin{cases} p & \text{if } d_i = \{dense\} \\ q & \text{if } d_i = \{sparse\} \end{cases} \tag{1}$$

Timing, Actions and Payoff

First, agents simultaneously vote for their preferred monitor $m_i \in \{0, 1\}$, where m_i=0 implies no monitor is chosen by individual i and m_i=1 means i votes for having the monitor. Once the participants cast their votes, a monitoring technology is assigned to the group according to the following rule

$$m^* = \begin{cases} 1 & \text{if } m_i = m_j = 1 \\ 0 & \text{if otherwise} \end{cases} \tag{2}$$

where m^* denotes the outcome of the vote. Second, agents make their contribution decision $c_i \in \mathbb{R}$. The action profile of agent i is then $(m_i,\ c_i)$. The total contribution is increased by 50 % and divided equally among the group meaning the rate of return for the contribution game with two players is $\frac{3}{4}$.

The utility of Player i is a function of both c_i and c_j, of the level of altruism α_i and the rate of return in the contribution game $\frac{3}{4}$. We assume a convex cost c^2 in order to ensure there is an interior solution. We also believe this functional form characterized by increasing marginal cost well represents the behavioral burden of contributing. Further, in the spirit of psychological games[19], how much player i values the utility of player j depends on $i's$ belief about the altruism of player j, $\mu_{0i}(\alpha_j)$, since we believe that in this context belief dependant motivations deeply affect player $i's$ actions. In this regard, we take inspiration by Rabin (1993), who models reciprocity where players wish to act kindly (unkindly) as a

---

[18] This assumption is in line with the literature on altruism where individuals contribute and cooperate more in closely knit dense groups. See Liedler et al. (2009) and Goeree et al. (2010).

[19] For a review on psychological game refer to Dufwenberg (2008) and Attanasiy and Nage (2008).

function of belief about the other's type. The payoff of player i in the contribution game without a monitor is

$$U(\alpha_i|m^* = 0) = W - c_i - c_i^2 + \frac{3}{4}(c_i + c_j) + \alpha_i \cdot \mu_{0i}[W - c_j - c_j^2 + \frac{3}{4}(c_i + c_j)] \qquad (3)$$

In the case where a monitor is elected, we add two terms to the above utility function: a monitor cost mc and a reputation cost -$\delta P(c_i < \theta)$. Voting for the monitor is costly and i pays mc if she votes for the monitor irrespective if the monitor gets elected or not. If elected, the monitor can impose a reputation cost on the players. We introduce a parameter $\delta > 1$ which represents the penalty from a contribution lower than the social norm θ in the presence of the monitor. For the sake of exposition, we use a fixed value of δ. However, we could incorporate varying power of the monitor depending on their centrality by allowing $\delta \; \epsilon \; \{\delta_H, \; \delta_L\}$, where $\delta_H > \delta_L$ i.e. high central monitors are more effective in spreading information. The social norm is a stochastic parameter given that different monitors would have a different view of what a socially acceptable norm is. It is assumed to be uniformly distributed between $[0, \, \bar{\theta}]$ where $\bar{\theta}$ is the highest possible contribution. It can also be interpreted as a reference point (Kahneman and Tversky 1991) that varies with each monitor, i.e. it hinges on the distribution of θ. The probability of one's contribution being higher than the norm is then simply the cumulative distribution of all contributions which is $\frac{c_i}{\theta}$ and the probability of contributing below the acceptable social - as perceived by the monitor - norm can thus be represented as

$$P(c_i < \theta) = \begin{cases} 1 - \frac{c_i}{\theta} & \text{if } c_i < \theta \\ 0 & \text{if otherwise} \end{cases} \qquad (4)$$

The utility if a monitor is elected ($m^* = 1$) is then $U_i(\alpha_i|m^* = 1)=$

$$W - \hat{c}_i - {\hat{c}_i}^2 + \frac{3}{4}(\hat{c}_i + \hat{c}_j) - mc - \underbrace{\delta P(\hat{c}_i < \theta)} + \underbrace{\alpha_i \cdot \mu_{0i}[W - \hat{c}_j - \lambda {\hat{c}_j}^2 + \frac{3}{4}(\hat{c}_i + \hat{c}_j) - mc - \delta P(\hat{c}_j < \theta]} \qquad (5)$$

Moreover, players are Bayesian and i updates her prior about $j's$ type $\mu_{0i}(\alpha_j)$ via Bayesian rule to $\mu_{1i}(\alpha_j)$ depending on the outcome of the voting , m^*.

Equilibrium

We assume that the altruism parameter α_i of individual i fully determines her demand for peer monitoring. More formally, we consider an equilibrium of the form below. An (altruistic) player i of type α_h cares very much about the utility of the other player irrespective of $j's$ type. She would therefore prefer not to elect a monitor [20] in order to avoid the other player being punished via bad reputation in case of low contribution. For a player i of type α_l, however, the cost of electing a monitor and the negative reputation effects for both herself and j is outweighed by the increase in group contribution driven by the presence of the monitor. Agents then contribute differently depending on their type α_i, outcome of the vote m^* and updated belief μ_{1i} about player j, once the outcome of the vote is revealed. The separating equilibrium would then be

$$\sigma_i(\alpha_i) = \begin{cases} m_i = 0 & \text{if } \alpha_i = \alpha_h \\ m_i = 1 & \text{if } \alpha_i = \alpha_l \end{cases} \tag{6}$$

Given the equilibrium above, when $\alpha_i = \alpha_l$, i would always vote for a monitor and, given the voting rule m^* defined above, she is able to perfectly infer the voting choice of player j. In this case, i updates her prior to $\mu_{1i}(\alpha_j) = 1$ if $m^* = 0$ and $\mu_{1i}(\alpha_j) = 0$ if $m^* = 1$. On the other hand, type α_h always votes for $m_i = 0$ and no monitor is elected - $m^* = 0$ - irrespective of the vote of the other player. In this case player i cannot infer anything about j's type and she sticks to the prior $\mu_{1i}(\alpha_j) = \mu_{0i}(\alpha_j)$. First, we solve the above set of equations and calculate the value of optimal contributions across the different scenarios. Secondly, given c_i, we study when the above separating equilibrium holds true.

We find that for type α_l, voting for the monitor is an optimal strategy for certain values of initial prior $\mu_{0i}(\alpha_j) < p^*_{\alpha_l}$. On the other hand, for type α_h it is always a dominant strategy to vote for no monitor. Therefore in order for there to be a separating equilibrium, it should be the case that people have a low prior on the proportion of altruists.

Proposition 1. *In the game with exogenous monitors, high-type players α_h contribute always more than low-type ones α_l. Moreover, at equilibrium the contributions of both players' types are higher in the presence of the monitor than without,*

$$\hat{c}_i^{exo} > c_i^{exo}$$

for $i = h, l$

---

[20] Given that dense groups have higher subjective probability of being altruists, the demand of peer monitoring should be lower than that in sparse as seen in Fig 3.

where $\hat{c}_i^{exo}$ indicates the optimal contribution when the contribution game is played in presence of a monitor and c_i^{exo}. The result is simply driven by the reputation effect of the monitor, which can entail the penalty δ in case of contributions lower than the social norm θ.

We now study the optimal contributions in the setting with the endogenous election of the monitor. In this case, the elected monitor serves as a signal of each other's types. Players are Bayesian and update their prior beliefs about the opponent's type, knowing their own vote in the election stage. The subjects play the game twice in random order, once with the endogenous monitor and once with the exogenous. When they play with the endogenously elected monitor first, they update their beliefs before the first contribution and we assume they will contribute identical amounts in the rounds, given their constant beliefs across rounds. On the other hand, when they play first with the exogenously assigned monitor, the updating happens only at the second round.

Proposition 2. *When the endogenous selection of monitors is allowed, the average contribution across the exogenous and the endogenous selection of monitors is higher if endogenous is played first and no monitor ($m^* = 0$) is elected by the group. If the group elects a monitor ($m^* = 1$), contribution decreases if endogenous is played first only in dense group.*

The election of no monitor signals that the group peers are high types α_h, thus pushing the posterior belief up to $\mu_{1i} = 1$. When this happens in the first round, contribution in both rounds increases since updating of the prior happens before players choose their optimal contribution. Surely, in this case it is higher then when the updating occurs only in the second round. On the other hand, when a monitor is elected in the first round, contribution decreases in both rounds. The decrease is however starker for dense groups due to higher initial prior. Hence, this effect holds only in dense groups.

Proposition 3. *Let us assume that δ is large enough and θ is small enough. Then, there exists a value of the initial prior $p^*_{\alpha_l}$ such that for $p < p^*_{\alpha_l}$, low types players α_l are better off voting for no monitor. Moreover, for $0 < p < p^*_{\alpha_l}$, there always exists the separating equilibrium σ, where low types α_l vote for the monitor and high types α_h vote for no monitor.*

We believe that the assumption of large δ is quite natural, given that in our context formal institutions are weak, and reputation concerns drive most of the social interactions. Similarly, the ex-ante level of cooperative behavior of these villages is modest, hence justifying the assumption of low values of θ. The mechanism underlying this proposition lies in the fact that high type players α_h always vote for no monitor, irrespective of the group they are in. Moreover, this proposition gives us reason to believe that a story of reciprocal altruism where players strategically vote for peer monitoring and contribute to a local public good, well describes the voting behavior we see in the experimental data, i.e. players vote more often for having a monitor in sparse groups (low p) rather than in dense ones.

5 Conclusion

By using original network data and a novel design, we try to understand how the varying demand of peer monitoring depends on group density and how this in turn affects cooperation. We divide the network into groups of three individuals with varying network distance, where dense implies each individual is at most at distance 2 (average path length < 1.6) and sparse implies each individual is at least at distance 4 (average path length > 4). To begin with, we show that dense groups prefer to not have a monitor whereas sparse groups choose to have a central one, reflecting variation in trust. Low central monitors are seldom chosen. In line with previous literature, when individuals are socially close (dense), they can sustain a higher level of cooperation without outside intervention. Dense groups contribute higher than the sparse group in the contribution game.

Next, we show 'how' an institution is assigned matters for cooperation. The endogenous choice of monitoring increases cooperation in only sparse groups. Looking at the order of the monitor treatment, the outcome of the vote being revealed in endogenous treatment carries an additional information regarding individual preferences and hence, when revealed, acts as a signal to the group. When endogenous treatment is played first and no monitor is chosen by the group, individuals tend to contribute higher in both groups. However, when endogenous treatment is played first and a monitor is chosen, contribution decreases only in dense groups due to a stronger prior about the level of altruism. This is an interesting finding that suggests monitoring should be catered to the needs of the community. It is also in line with the argument that repeated interactions in dense groups imply higher concern for reputation.

Given the increased popularity of community-based interventions and focus on peer monitoring, it is important to understand the role social networks play in small scale societies. We propose here a theoretical framework followed by a simple experiment that show that the effect of a monitor can be very different depending on the density of the network. Our work opens avenues for further research. We would like to understand the choice of the monitors further by presenting individuals with a panel of monitor options rather than just the high and low central ones.

References

[1] Andreoni, J. and Sanchez, A. (2020). Fooling myself or fooling observers? avoiding social pressures by manipulating perceptions of deservingness of others. *Economic Inquiry*, 58.

[2] Attanasiy, G. and Nagelz, R. (2008). Psychological games. In *Games, Rationality and Behaviour. Essays on Behavioural Game Theory and Experiments*. Palgrave Macmillan, Houndmills.

[3] Baldassarri, D. and Grossman, G. (2013). The effect of group attachment and social position on prosocial behavior. evidence from lab-in-the-field experiments. *PLOS ONE*, 8(3).

[4] Ballester, C., Calvo-Armengol, A., and Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, 74(5).

[5] Banerjee, A., Duflo, E., Jackson, M. O., and Chandrashekhar, A. (2013). The diffusion of microfinance. *Science*, 341(6144).

[6] Banerjee, A., Duflo, E., Jackson, M. O., and Chandrashekhar, A. (2019). Using gossips to spread information: Theory and evidence from two randomized controlled trials. *The Review of Economic Studies*.

[7] Beersma, B. and Van Kleef, G. A. (2011). How the grapevine keeps you in line: Gossip increases contributions to the group. *Social Psychology and Personality Science*, 2.

[8] Bjorkman, M. and Svensson, J. (2009). Power to the people: Evidence from a randomized field experiment on community-based monitoring in uganda. *The Quarterly Journal of Economics*, 124(2).

[9] Bowles, S. (2008). Economic experiments undermine "the moral sentiments": Evidence from policies designed for self-interested citizens. *Science*, 320(1605).

[10] Breza, E. and Chandrashekhar, A. (2018). Social networks, reputation and commitment: evidence from a savings monitors experiment. *Econometrica*.

[11] Breza, E., Chandrashekhar, A., and Larreguy, H. (2016). Network centrality and informal institutions: Evidence from a lab experiment in the field. *NBER working paper 20309*.

[12] Chandrashekhar, A., Kinnan, C., and Larreguy, H. (2018). Social networks as contract enforcement: Evidence from a lab experiment in the field. *American Economic Journal: Applied Economics*.

[13] Charness, Gary Cobo Reyes, R. and Jimenez, N. (2008). An investment game with third-party intervention. *Journal of Economic Behavior and Organization*, 68.

[14] Dal Bo, Pedro Foster, A. and Putterman, L. (2010). Institutions and behavior:experimental evidence on the effects of democracy. *American Economic Review*, 100.

[15] Dufwenberg, M. (2008). Psychological games. In *Palgrave Macmillan (eds) The New Palgrave Dictionary of Economics*. Palgrave Macmillan, London.

[16] Etang, Alvin Fielding, D. and Knowles, S. (2011). Does trust extend beyond the village? experimental trust and social distance in cameroon. *Experimental Economics*, 14(1).

[17] Fehr, D. and Sutter, M. (2019). Gossip and the efficiency of interactions. *Games and Economic Behavior*, 113.

[18] Fehr, E. and Gachter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4).

[19] Ferraro, P. J. and Vosslery, C. A. (2010). The source and significance of confusion in public goods experiments. *The B.E. Journal of Economic Analysis Policy Contributions*, 10(1).

[20] Fonseca, M. A. and Peters, K. (2018). Will any gossip do? gossip does not need to be perfectly accurate to promote trust. *Games and Economic Behavior*, 107.

[21] Galbiati, Roberto Vertova, P. (2008). Obligations and cooperative behaviour in public good games. *Games and Economic Behavior*, 64:146–170.

[22] Gelade, W. and Guirkinger, C. (2018). The enforcement advantage of external monitoring: Lessons from an experiment with joint-liability groups in burkina faso:. *Journal of Economic Behavior and Organization*, 151.

[23] Glaeser, E., Laibson, D., Scheinkman, J. A., and Soutter, C. (2000). Measuring trust. *Quarterly Journal of Economics*, 115(3).

[24] Glockner, A., Kube, S., and Nicklisch, A. (2018). The joint benefits of observed and unobserved social sanctions. *Journal of Behavioral and Experimental Economics*, 75.

[25] Goeree, J., McConnell, M., Mitchell, T. D., Tromp, T., and Yariv, L. (2007). The 1/d law of giving. *American Economic Journal: Microeconomics*, 2(1):183–203.

[26] Greif, A. (1989). Reputation and coalitions in medieval trade: Evidence on the maghribi traders. *Journal of Economic History*, 49(4).

[27] Greif, A. (2006). *Institutions and the path to the modern economy: Lessons from medieval trade*. Cambridge University press.

[28] Grossman, G. and Baldassarri, D. (2012). The impact of elections on cooperation: Evidence from a lab-in-the-field experiment in uganda. *American Journal of Political Science*, 56(4).

[29] Herold, F. (2010). Contractual incompleteness as a signal of trust. *Games and Economic Behavior*, 68.

[30] Hoffman, Elizabeth McCabe, K. and Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review*, 3(86).

[31] Jackson, M. O., Tomas, R. B., and Tan, X. (2012). Social capital and social quilts: Network patterns of favor exchange. *American Economic Review*, 102(5).

[32] Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2).

[33] Kandori, M. (1992). Social norms and community enforcement. *The Review of Economic Studies*, 59.

[34] Karlan, D., Mobius, M., Rosenblat, T., and Szeidl, A. (2009). Trust and social collateral. *The Quarterly Journal of Economics*, 124(3).

[35] Kosfeld, M. and Rustagi, D. (2015). Leader punishment and cooperation in groups: Experimental field evidence from commons management in ethiopia. *American Economic Review*, 105(2).

[36] Kranton, R. E. (1996). Reciprocal exchange: A self-sustaining system. *The American Economic Review*, 86(4).

[37] Leider, S., Mobius, M. M., Rosenblat, T., and Quoc-Anh, D. (2009). Directed altruism and enforced reciprocity in social networks. *Quarterly Journal of Economics*, 124(4).

[38] Mansuri, G. and Rao, V. (2004). Community-based and -driven development: A critical review. *World Bank Policy Research Working Paper 3209.*

[39] Narayan, D. (1995). Designing community-based development. *Environmentally and Socially Sustainable Development Network, Washington, World Bank*, 108(1).

[40] Olken, B. A. (2007). Monitoring corruption: Evdence from a field experiment in indonesia. *Journal of Political Economy*, 115(2).

[41] Ostrom, E. (1990). *Governing the Commons: The evolution of institutions for collective action*. Cambridge University press.

[42] Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economics Review*, 83(5).

[43] Sommerfeld, R., Krambeck, H.-J., Semmann, D., and Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. *Proceedings of the National Academy of Sciences of the U.S.A*, 104.

[44] Sutter, M., Haigner, S., and Kocher, M. G. (2010). Choosing the carrot or the stick? endogenous institutional choice in social dilemma situations. *The Review of Economic Studies*, 77(4).

[45] Tyran, J. and Feld, L. P. (2006). Achieving compliance when legal sanctions are non deterrent. *The Scandinavian Journal of Economics*, 108(1).

[46] Wu, J., Balliet, D., and Van Lange, P. (2015). When does gossip promote generosity? indirect reciprocity under the shadow of the future. *Social Psychology and Personality Science*.

[47] Wu, J., Balliet, D., and Van-Lange, P. (2016). Gossip versus punishment: The efficiency of reputation to promote and maintain cooperation. *Scientific reports*.

[48] Zelditch, M. (2001). Processes of legitimation: Recent developments and new directions. *Social Psychology Quarterly*, 64.

Figures

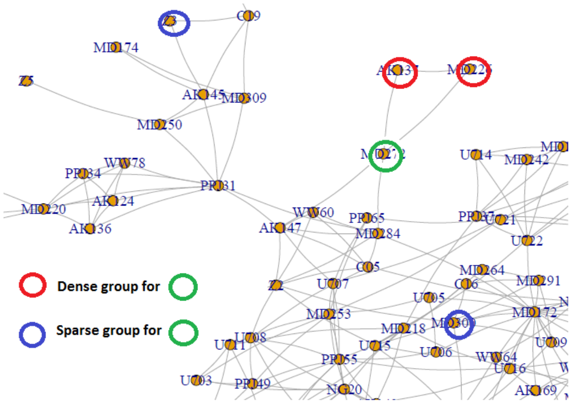


Figure 1 Example of formation of groups

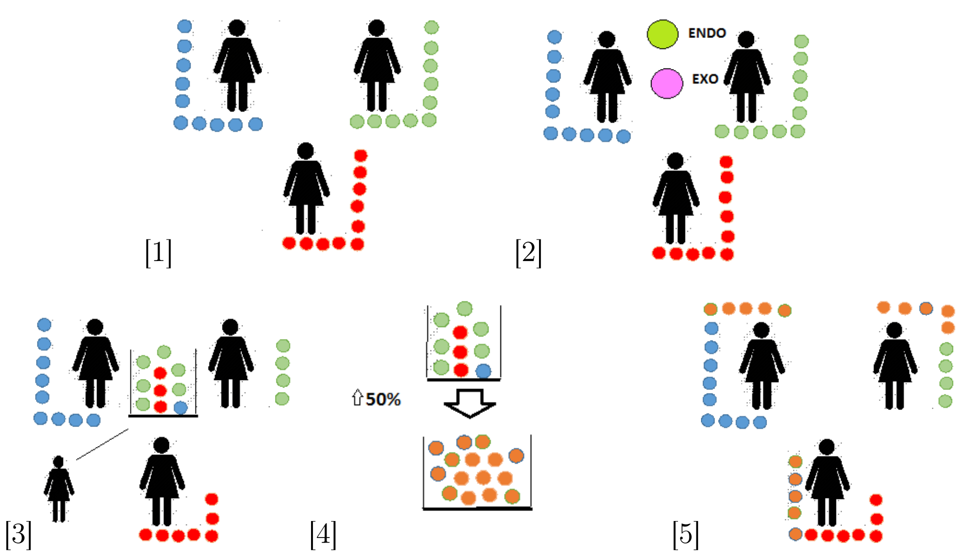


Figure 2 Experimental Design

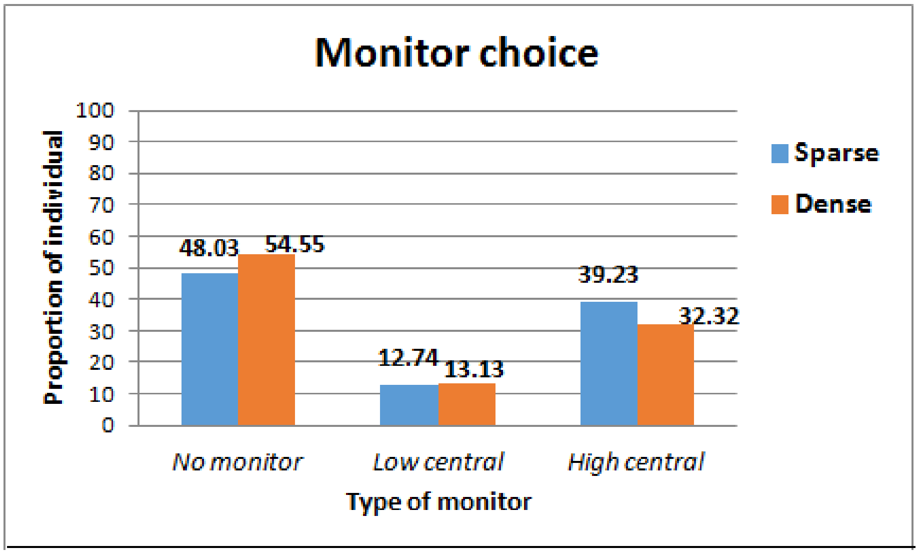


Figure 3 Percentage of individuals voting in Sparse and Dense groups

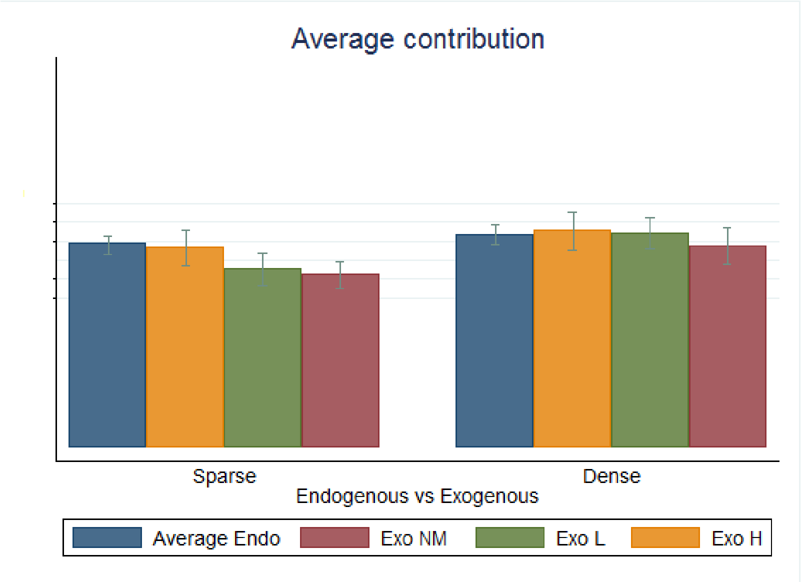
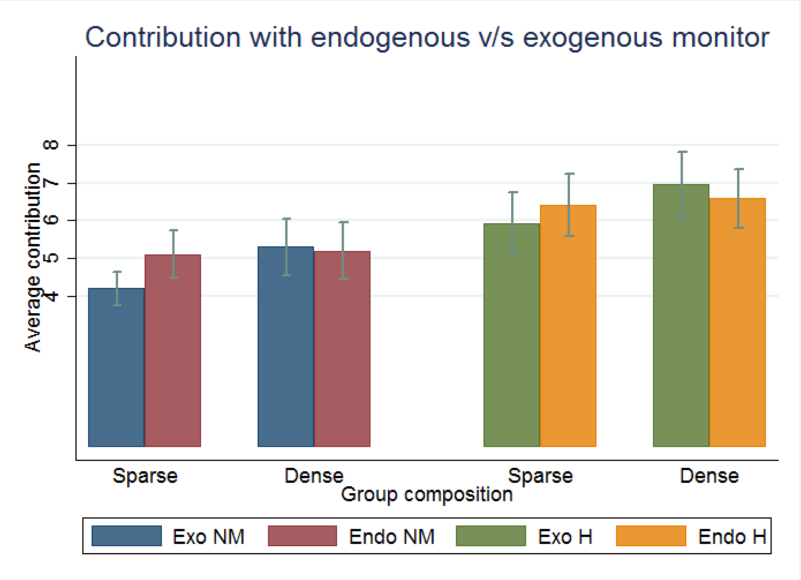
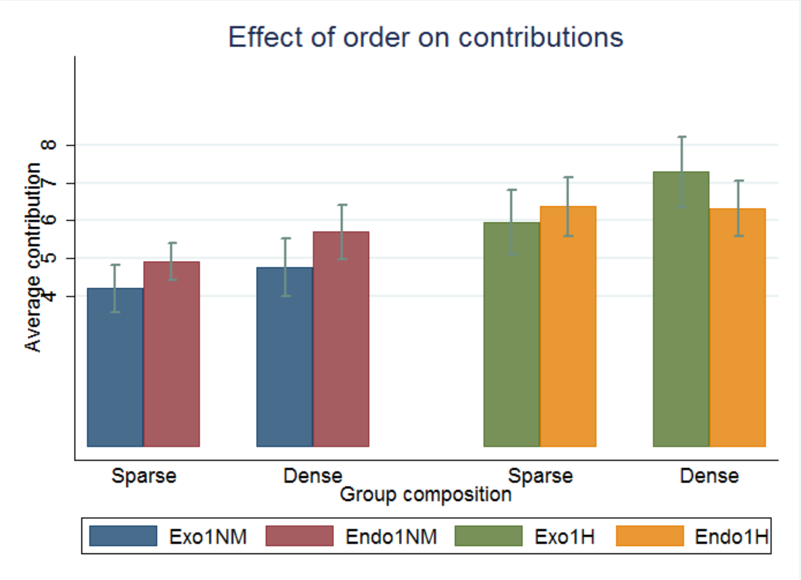


Figure 4 Average contribution endogenous v/s exogenous monitors with selection



In the bar graph, y-axis represents group composition and x-axis represents average contribution. We focus on a sub sample where the same group plays under the same monitoring condition both exogenously and endogenously. The blue bar represents no monitor being assigned exogenously as compared to the red bar where monitor is assigned endogenously. The green bar represents exogenously assigned high central monitor as compared to the yellow bar.

Figure 5 Contribution with endogenous v/s exogenous monitors without selection



In the bar graph, y-axis represents group composition and x-axis represents average contribution. We focus on a sub sample where the same group plays under the same monitoring condition both exogenously and endogenously. The blue bar represents no monitor being assigned exogenously as compared to the red bar where monitor is assigned endogenously. The green bar represents exogenously assigned high central monitor as compared to the yellow bar.

Figure 6 Order of endogenous and exogenous monitors

Tables

		Dense group		
		No monitor	*Low central*	*High central*
Sparse group	*No monitor*	34.95%	4.57%	9.34%
	Low central	5.17%	4.37%	2.98%
	High central	14.71%	4.17%	19.68%

Table 1 Variation in voting within individual across different groups

	NM	**L**	**H**
DENSE	5.39	5.71	5.84
SPARSE	4.67	4.76	5.41

Notes: Dense group contribute more than the sparse ones. In the presence of a high central monitor, contribution increases significantly in sparse groups.

Table 2 Average contribution in the exogenous treatment

	Monitor choice	Monitor choice
Low central		
Dense	-0.062	0.062
	(0.24)	(0.20)
High central		
Dense	-0.407***	-0.466**
	(0.18)	(0.20)
N	503	459
Group characteristics	No	Yes

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: No monitor is the base outcome. Monitor choice refers to the individual choice out of: No monitor, High central monitor and Low central monitor. Elected monitor is choice at the level of the group. Dense is a dummy variable that takes value 1 if the group is dense (average path length <2) and 0 otherwise. We control for individual and round fixed effects.

Table 3 Multilogit regression on monitor choice

	contribution	contribution
Dense	1.407***	1.614***
	(0.31)	(0.33)
H	0.760**	0.939**
	(0.36)	(0.38)
H × Dense	-0.933**	-1.088**
	(0.43)	(0.44)
L	0.646*	0.748*
	(0.38)	(0.39)
L × Dense	-0.972**	-1.170***
	(0.42)	(0.45)
N	503	459
Group characteristics	No	Yes

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: Dense is a dummy variable that takes value 1 if the group is dense (average path length <2) and 0 otherwise. H is a dummy variable which is 1 if a High central monitor is elected and L is a dummy variable which is 1 if a Low central monitor is elected. We control for individual and round fixed effects along with individual choice.

Table 4 Fixed effect regression on average contribution under exogenous monitors

	Contribution			
	Sparse(H)	Dense(H)	Sparse(NM)	Dense(NM)
Endogenous	0.50**	-0.377	0.913***	-0.111
	(0.24)	(0.25)	(0.22)	(0.20)
N	112	106	184	144

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: Contribution is the amount given by individuals under each sub group. Sparse (H) refers to sparse groups (average path length > 4) who played both endogenous and exogenous treatment under a High central monitor (H). Dense NM refers to dense groups (average path length < 2) who played both endogenous and exogenous treatment under no monitor. Endogenous is a dummy that takes value 1 if contribution was made with choice of the group. We control for individual and round fixed effects.

Table 5 Endogenous v/s Exogenous Contribution Regression FE

	Contribution			
	Sparse(H)	Dense(H)	Sparse(NM)	Dense(NM)
Order	0.281	-1.368**	0.770*	1.273**
	(0.58)	(0.59)	(0.42)	(0.64)
N	92	94	158	120

* p<0.10, ** p<0.05, *** p<0.01

Notes: Contribution is the amount given by individuals under each sub group. Sparse (H) refers to sparse groups (average path length > 4) who played both endogenous and exogenous treatment under a High central monitor (H). Dense NM refers to dense groups (average path length < 2) who played both endogenous and exogenous treatment under no monitor. Order is a dummy that takes value 1 if endogenous treatment was played first. We control for individual characteristics and have round fixed effects.

Table 6 Effect of order on Contribution

Appendix

5.1 Appendix A

Important clarification

The text in italic is not meant to be read aloud to experiment participants. It has the explanation of what experimenters should do. The remaining text that is not in italics is meant to be read aloud to experiment participants.

Experiment

Divide the research team into two groups: team A and team B. As participants enter the venue, team A must welcome them and locate their ID number based on their name from the individual identification list. The research team must then provide the participants with the consent forms, read the forms aloud, explain to them the contents of the forms and that the participants are free to leave at their discretion, answer any questions participants may have, and obtain their consent. [Go to Consent Form]

Then, team B should be ready to enter data on contributions.

Experiment begins

Thanks for coming today! We are researchers from Rooster Logic. You are participating in a study on daily decision-making. Today you will play a series of short games. The information gathered here will be confidential and used for research purposes only.

Overview

Today, we will ask you to play a game with two different groups of people for two rounds each. You will randomly be placed in groups of three for the game, whose identity will be known. In each game, you and your group members will make some decisions. The result of these decisions will determine how much money you will earn today.

The games will represent situations and decisions you make every day in your life. You earn some money, you keep some money for yourself, you might give some money to your neighbors or friend, use the money to fund a common project etc.

Explanation of payment

Let us now discuss how you will make money today. First, you will receive 100 Rs. for simply participating in our games. Second, you will make money from the decisions made during the game.

You will play the same game with two different groups. In the beginning of each game, you will get some income in the form of tokens in a bag we call an 'INCOME POT'. The game is easy and all that you need to do is decide how many tokens you want to keep for yourself and how many tokens you want to contribute to the 'PUBLIC POT'. The total amount collected in the 'PUBLIC POT' will be increased in value by 50%. In both games, the experimenter will collect the tokens that you want to contribute in two different 'PUBLIC POT'.

At the end of the experiment, we will pick one 'PUBLIC POT' out of the 4 and the total amount with the additional 50% increase will be equally divided among the four players in your corresponding group. You will receive equal share, irrespective of how much you put in the 'PUBLIC POT', Respectively, the tokens you decided to keep for yourself in the 'INCOME POT' corresponding to that game will be yours.

Demonstrate: The "experimenter" should explain that they will be playing four rounds during the day with two different groups of people. Please show them the graphical image and explain how the contribution game works and how they would earn.

See then that the decisions you make in all rounds count but you will only be paid the amount in one randomly chosen game. Before I explain the game you will play today in detail, are there any questions?

Answer any questions that they may have.

Explanation of the game

The game I will explain to you is a very simple one. In this game, you will be matched randomly with 3 more people who you will interact with. You are not allowed to talk to each other throughout this game. At the beginning of the game, you and your partners will get some money that you can either keep for yourself or contribute to a common pot.

There are two stages in this game: First you will be given the choice to elect a monitor to oversee the contribution game that we just briefly explained. The monitor vote will be followed by the contribution task. Let me explain in detail what the contribution task is.

At the beginning of each game, each of you will be given an initial income of Rs 100. All earnings during the games will be represented by tokens, each with a value of Rs 10. Then, each of you will be provided with 10 tokens that are worth Rs 100 in total. This cup will be known as 'INCOME POT'.

Demonstrate procedure, the objective you should have in mind is that individuals acquire a sense of the physicality of the game.

Now, we will explain how you can use your income in the game. You can either keep the tokens for yourself in the 'INCOME POT' or you can contribute to the 'PUBLIC POT'. The money that you decided to keep in the 'INCOME POT' will be yours. The tokens that you will put in the 'PUBLIC POT' will be added to the tokens that rest of your group put in the 'PUBLIC POT'. The total amount contributed by the group will then increase in value by 50%.

The amount you contribute to the 'PUBLIC POT' will not be revealed to the rest of the members of your group. To contribute to the 'PUBLIC POT', you will give the number of tokens you want to contribute to the experimenter in the 'PUBLIC POT'. Remember that 1 token is worth 10 Rs.

Demonstrate the procedure via the chart again. Explain to them that 2 tokens= 20 Rs

In the first stage, you will be given a chance to elect a monitor to oversee this contribution task. The monitor will observe the amount contributed by each individual to the 'PUBLIC POT' which is otherwise not known. In order to choose a monitor, you will put a tick next to one of the two choices: either having a monitor or not having a monitor. If you decided to have a monitor by putting a tick on the square, you will choose the name of the person you want to elect in the same sheet. If you decide to vote for having a monitor, you will be charged 10 Rs from the money you have been given for participation in the game.

Demonstrate the voting sheet to participants.

We will consider the choices of everyone in your group. The option that gets the highest number of votes will be chosen. Now, to see whether the majority choice will be implemented or an external option will be randomly assigned, we will pick a ball from this box without looking. In the box which we will call the 'CHOICE BOX'.

We have two balls, 1 Pink and the other Green. We will pick a ball from the box, if a green ball is chosen, then the option chosen by the group will be implemented. If a pink ball is chosen instead, we will randomly assign one of the 3 options to your group.

Demonstrate the voting procedure to the participants with four enumerators. Make sure they understand the use of the 'CHOICE BOX'

Do we have any questions at this point? Have you understood the two stages of the game? Now, we will demonstrate the complete game.

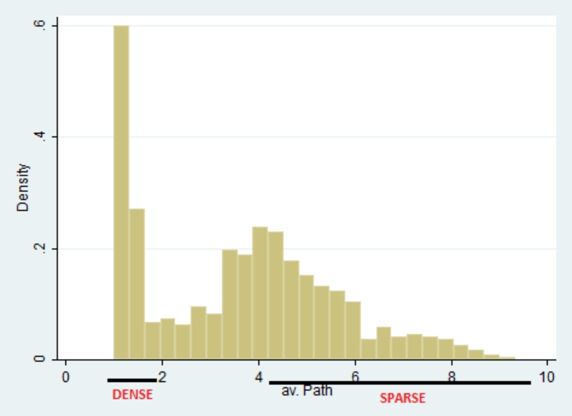
Five members of the team of experimenters should do the demonstration. Four of them should take the role contributors. The fifth person should represent himself and we will refer to him/her as the "experimenter."

Do you have any questions?

Now, we will practice the game. Note that this will only be practice rounds and that you will not actually play with your actual partner. You will play the actual games with your actual partners after we explain the contribution game, practice them and we answer any question you might have about the games.

Participants play three rounds of the game and information is recorded exactly as if the game was actually being played.

5.2 Appendix B



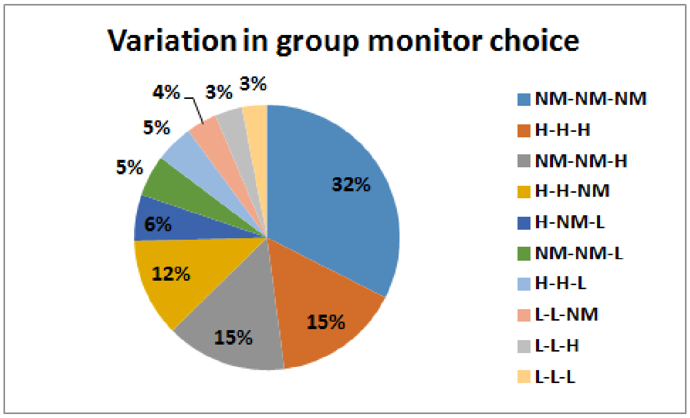
This is the distribution of average path length in the 1006 groups we formed. Average path length is defined as the average number of steps along the shortest paths for all possible pairs of the group. We over sampled closely knit groups with average path length <2 (dense). Sparse group is defined as groups with average path length >4.

Figure 7 Distribution of average path length across groups

	Mean	**Std.dev**	**N**
Individual Characteristics			
age	35.8	11.43	503
education	3.06	3.85	503
no, of links	11.38	4.46	503
centrality	0.052	0.071	503
wealth index	-0.253	1.503	503
Group Characteristics			
Same caste	0.74	0.438	503
Same education	0.3801	0.485	503

Table 7 Summary stats

5.3 Appendix C: Monitor choice



It shows the variation in individual choice within a group. NM: no monitor being chosen, L: low central monitor and H: high central monitor is chosen. In most groups, all three members vote for NM followed by all three group members voting for H.

Figure 8 Variation in individual choice within a group

	Monitor	
	Dense	Sparse
Age	-0.0004	-0.003*
	(0.002)	(0.001)
Caste	0.021	0.090***
	(0.02)	(0.02)
Education	-0.030**	-0.052***
	(0.01)	(0.01)
Wealth	0.014	0.050***
	(0.01)	(0.01)
Favor return strangers	0.002	-0.048***
	(0.01)	(0.01)
Help friends	-0.009	0.038***
	(0.01)	(0.01)
Centrality	0.181	0.566**
	(0.30)	(0.29)
Distance to H	-0.001	0.032***
	(0.01)	(0.01)
Distance to L	0.005	0.010*
	(0.01)	(0.01)
N	842	842

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes:Monitor is a dummy that takes value 0 if no monitor is elected and 1 if either a high or low monitor is elected. The first column (Dense) regresses individual characteristics with outcome of the vote and the second column does the same but for sparse groups.

Table 8 OLS regression for monitor choice behavior

5.4 Appendix D: Model

Let us first consider the exogenous case with no signalling. Since the monitoring technology is randomly assigned and not chosen by the group, there is no update of the prior μ_{0i}.

For type α_h

$$U(\alpha_i = \alpha_h, \cdot, m^* = 1) = p \cdot U(\hat{c_h}^{exo}, \hat{c_h}^{exo}, \alpha_h, \alpha_h) + (1-p) \cdot U(\hat{c_h}^{exo}, \hat{c_l}^{exo}, \alpha_h, \alpha_l)$$

$$U(\alpha_i = \alpha_h, \cdot, m^* = 0) = p \cdot U({c_h}^{exo}, c_h^{exo}, \alpha_h, \alpha_h) + (1-p) \cdot U({c_h}^{exo}, c_l^{exo}, \alpha_h, \alpha_l)$$

For type α_l

$$\begin{aligned} U(\alpha_i = \alpha_l, \cdot, m^* = 0) &= p \cdot U(c_l^{exo}, c_h^{exo}, \alpha_l, \alpha_h) + (1-p) \cdot U(\hat{c_l}^{exo}, \hat{c_l}^{exo}, \alpha_l, \alpha_l) \\ U(\alpha_i = \alpha_l, \cdot, m^* = 1) &= p \cdot U(c_l^{exo}, c_h^{exo}, \alpha_l, \alpha_h) + (1-p) \cdot U(\hat{c_l}^{exo}, \hat{c_l}^{exo}, \alpha_l, \alpha_l) \end{aligned}$$

where $\hat{c}_i^{exo}$ denotes the contribution of player i when there is the monitor and c_i^{exo} when there is no monitor.

Solving for each contribution level c_l^{exo}, c_h^{exo}, $\hat{c_l}^{exo}$, $\hat{c_h}^{exo}$ we get for the exogenous assignment of monitoring technology,

$$\begin{aligned} \hat{c_l}^{exo} &= \frac{3\alpha_l p - 1}{8} + \frac{\delta}{2\theta} & \qquad c_l^{exo} &= \frac{3\alpha_l p - 1}{8} \\ \hat{c_h}^{exo} &= \frac{3\alpha_h p - 1}{8} + \frac{\delta}{2\theta} & \qquad c_h^{exo} &= \frac{3\alpha_h p - 1}{8} \end{aligned}$$

In the endogenous case we have to take into account the election rule and now the monitor outcome (m^*) becomes a signal according to which players update their belief about player j's type. We show that in both dense and sparse groups, average contribution in two rounds is higher if endogenous is played first and no monitor ($m^* = 0$) is elected by the group. In the event of a monitor ($m^* = 1$) being elected by the group, contribution decreases significantly only in dense groups. Given the updated priors, we can write the utility function for type α_h

$$\begin{aligned} U(\alpha_i = \alpha_h, \cdot, m^* = 1) &= U(\hat{c_h}^{end}, \hat{c_l^{end}}, \alpha_h, \alpha_l) \\ U(\alpha_i = \alpha_h, \cdot, m^* = 0) &= p \cdot U(c_h^{end}, c_h^{end}, \alpha_h, \alpha_h) + (1-p) \cdot U(c_h^{end}, c_l^{end}, \alpha_h, \alpha_l) \end{aligned}$$

For type α_l

$$\begin{aligned} U(\alpha_i = \alpha_l, \cdot, m^* = 0) &= U(c_l^{end}, c_h^{end}, \alpha_l, \alpha_h) \\ U(\alpha_i = \alpha_l, \cdot, m^* = 1) &= U(\hat{c_l^{end}}, \hat{c_l^{end}}, \alpha_l, \alpha_l) \end{aligned}$$

Solving for each contribution level c_l^{end}, c_h^{end}, $\hat{c_l}^{end}$, $\hat{c_h}^{end}$ we get,

$$\hat{c_l^{end}} = -\frac{1}{8} + \frac{\delta}{2\theta} \qquad c_l^{end} = \frac{3\alpha_l - 1}{8}$$
$$\hat{c_h}^{end} = -\frac{1}{8} + \frac{\delta}{2\theta} \qquad c_h^{end} = \frac{3\alpha_h p - 1}{8}$$

To establish the separating equilibrium,

For type: $\alpha_i = \alpha_l$

To show that voting for a monitor is optimal we calculate ex-ante utility of player i using prior beliefs via ,

$$U(\alpha_i = \alpha_l | m_i = 1) = p \cdot U(\alpha_l, \alpha_h, c_l, c_h) + (1-p) \cdot U(\alpha_l, \alpha_l, \hat{c_l}, \hat{c_l})$$
$$U(\alpha_i = \alpha_l | m_i = 0) = p \cdot U(\alpha_l, \alpha_h, c_l, c_h) + (1-p) \cdot U(\alpha_l, \alpha_l, c_l, c_l)$$
$$U(\alpha_i = \alpha_l | m_i = 1) - U(\alpha_i = \alpha_l | m_i = 0) = (1-p) \cdot \{U(\alpha_l, \alpha_l, \hat{c_l}, \hat{c_l}) - U(\alpha_l, \alpha_l, c_l, c_l)\} - p \cdot mc$$

There exists a $p*_{\alpha_l}$ st. $p < p*_{\alpha_l}$ $U(\alpha_i = \alpha_l | m_i = 1) - U(\alpha_i = \alpha_l | m_i = 0)$ is positive.

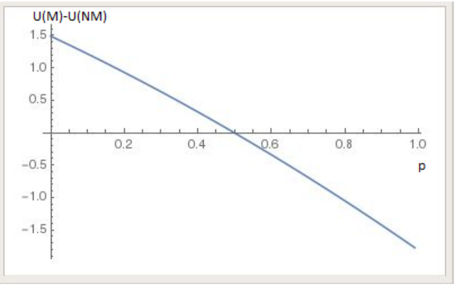


Figure 9 $U(m_i = 1) - U(m_i = 0)$

For type: $\alpha_i = \alpha_h$

$$U(\alpha_i = \alpha_h | m_i = 0) = p \cdot U(\alpha_h, \alpha_h, c_h, c_h) + (1-p) \cdot U(\alpha_h, \alpha_l, c_h, c_l)$$
$$U(\alpha_i = \alpha_h | m_i = 1) = p \cdot U(\alpha_h, \alpha_h, c_h, c_h) + (1-p) \cdot U(\alpha_l, \alpha_l, \hat{c_h}, \hat{c_l})$$
$$U(\alpha_i = \alpha_h | m_i = 1) - U(\alpha_i = \alpha_h | m_i = 0) = (1-p) \cdot \{U(\alpha_h, \alpha_l, \hat{c_h}, \hat{c_l}) - U(\alpha_h, \alpha_l, c_h, c_l)\} - p \cdot mc$$

In calculating the difference, we find that $U(\alpha_h | m_i = 1) - U(\alpha_h | m_i = 0) < 0$, therefore a high type would always vote for no monitor.

Combining the two results above, there exists an interval of p where type α_l would choose a monitor whereas type α_h would choose no monitor. Therefore there exists a separating equilibrium with p st. $0 < p < p*_{\alpha_l}$ where the above equilibrium holds true. If $p > p*_{\alpha_l}$ then there is no separating equilibrium and both types vote for no monitor.

5.5 Appendix E: Model extension with three agents

We expand the two agent model to three agents for it to be more representative of the interaction. Let us first consider the exogenous case with no signalling. Since the monitoring technology is randomly assigned and not chosen by the group, there is no update of the prior μ_{0i}. The voting rule m^* is slightly different with no tie possible.

For type α_h

$$\begin{aligned} U(m^*=1) &= p(1-p)\cdot[U(\hat{c_h}^{exo},\hat{c_h}^{exo},\hat{c_l}^{exo})+U(\hat{c_h}^{exo},\hat{c_l}^{exo},\hat{c_h}^{exo})]+ \\ &\quad p^2\cdot[U(\hat{c_h}^{exo},\hat{c_h}^{exo},\hat{c_h}^{exo})+U(\hat{c_h}^{exo},\hat{c_l}^{exo},\hat{c_l}^{exo})] \\ U(m^*=0) &= p(1-p)\cdot[U(\hat{c_h}^{exo},\hat{c_h}^{exo},\hat{c_l}^{exo})+U(\hat{c_h}^{exo},\hat{c_l}^{exo},\hat{c_h}^{exo})]+ \\ &\quad p^2\cdot[U(\hat{c_h}^{exo},\hat{c_h}^{exo},\hat{c_h}^{exo})+U(\hat{c_h}^{exo},\hat{c_l}^{exo},\hat{c_l}^{exo})] \end{aligned}$$

For type α_l

$$\begin{aligned} U(m^*=1) &= p(1-p)\cdot[U(\hat{c_l}^{exo},\hat{c_h}^{exo},\hat{c_l}^{exo})+U(\hat{c_l}^{exo},\hat{c_l}^{exo},\hat{c_h}^{exo})]+ \\ &\quad p^2\cdot[U(\hat{c_l}^{exo},\hat{c_h}^{exo},\hat{c_h}^{exo})+U(\hat{c_l}^{exo},\hat{c_l}^{exo},\hat{c_l}^{exo})] \\ U(m^*=0) &= p(1-p)\cdot[U(\hat{c_l}^{exo},\hat{c_h}^{exo},\hat{c_l}^{exo})+U(\hat{c_l}^{exo},\hat{c_l}^{exo},\hat{c_l}^{exo})]+ \\ &\quad p^2\cdot[U(\hat{c_l}^{exo},\hat{c_h}^{exo},\hat{c_h}^{exo})+U(\hat{c_l}^{exo},\hat{c_l}^{exo},\hat{c_l}^{exo})] \end{aligned}$$

where $\hat{c}_i^{exo}$ denotes the contribution of player i when there is the monitor and c_i^{exo} when there is no monitor. Solving for each contribution level c_l^{exo}, c_h^{exo}, $\hat{c_l}^{exo}$, $\hat{c_h}^{exo}$ we get for the exogenous assignment of monitoring technology,

$$\begin{aligned} \hat{c_l}^{exo} &= \frac{2\alpha_l p-1}{4}+\frac{\delta}{2\theta} & \qquad c_l^{exo} &= \frac{2\alpha_l p-1}{4} \\ \hat{c_h}^{exo} &= \frac{2\alpha_h p-1}{4}+\frac{\delta}{2\theta} & \qquad c_h^{exo} &= \frac{2\alpha_h p-1}{4} \end{aligned}$$

Given the calculated level of contribution under the exogenous monitor, we reiterate

Proposition 1 where

$$\hat{c}_i^{exo} > c_i^{exo}$$

for $i = h, l$

In the endogenous case we have to take into account the election rule and now the monitor outcome (m^*) becomes a signal according to which players update their belief about player j's type. We show in Proposition 2 that in both dense and sparse groups, average contribution in two rounds is higher if endogenous is played first and no monitor ($m^* = 0$) is elected by the group. In the event of a monitor ($m^* = 1$) being elected by the group, contribution decreases more in dense groups for a small p, where p $< \frac{2}{3}$. Given the updated priors, we can write the utility function for type α_h

$$U(\alpha_i = \alpha_h, \cdot, m^* = 1) = U(\hat{c_h}^{end}, c_l^{\hat{end}}, c_l^{end})$$
$$U(\alpha_i = \alpha_h, \cdot, m^* = 0) = p(1-p)[U({c_h}^{end}, c_h^{end}, {c_l}^{end}) + U({c_h}^{end}, c_l^{end}, {c_h}^{end})] + p^2 U({c_h}^{end}, c_h^{end}, {c_h}^{end})$$

For type α_l

$$U(\alpha_i = \alpha_l, \cdot, m^* = 0) = U(\hat{c_l}^{end}, \hat{c_h}^{end}, \hat{c_h}^{end})$$
$$U(\alpha_i = \alpha_h, \cdot, m^* = 0) = p(1-p)[U({c_l}^{end}, c_h^{end}, {c_l}^{end}) + U({c_h}^{end}, c_l^{end}, {c_h}^{end})] + p^2 U({c_l}^{end}, c_l^{end}, {c_l}^{end})$$

Solving for each contribution level c_l^{end}, c_h^{end}, $\hat{c_l}^{end}$, $\hat{c_h}^{end}$ we get,

$$\begin{aligned} c_l^{\hat{end}} &= \frac{4\alpha_l - 3}{12} + \frac{\delta}{2\theta} & \qquad c_l^{end} &= \frac{2\alpha_l - 1}{4} \\ \hat{c_h}^{end} &= -\frac{1}{4} + \frac{\delta}{2\theta} & \qquad c_h^{end} &= \frac{4\alpha_h - 3}{12} \end{aligned}$$

To establish the separating equilibrium,

For type: $\alpha_i = \alpha_l$

To show that voting for a monitor is optimal we calculate ex-ante utility of player i using prior beliefs via ,

$$
\begin{aligned}
U(m_i = 1) = p(1-p) \cdot [&U(\hat{c_l}^{exo}, \hat{c_h}^{exo}, \hat{c_l}^{exo}) + U(\hat{c_l}^{exo}, \hat{c_l}^{exo}, \hat{c_h}^{exo})] + \\
p^2 \cdot [&U({c_l}^{exo}, {c_h}^{exo}, {c_h}^{exo}) + U(\hat{c_l}^{exo}, \hat{c_l}^{exo}, \hat{c_l}^{exo})] \\
U(m_i = 0) = p(1-p) \cdot [&U({c_l}^{exo}, {c_h}^{exo}, {c_l}^{exo}) + U({c_l}^{exo}, {c_l}^{exo}, {c_h}^{exo})] + \\
p^2 \cdot [&U({c_l}^{exo}, {c_h}^{exo}, {c_h}^{exo}) + U(\hat{c_l}^{exo}, \hat{c_l}^{exo}, \hat{c_l}^{exo})]
\end{aligned}
$$

There exists a $p{*}_{\alpha_l}$ st. $p < p{*}_{\alpha_l}$ $U(\alpha_i = \alpha_l | m_i = 1) - U(\alpha_i = \alpha_l | m_i = 0)$ is positive.

For type: $\alpha_i = \alpha_h$

$$
\begin{aligned}
U(m_i = 1) = p(1-p) \cdot [&U(\hat{c_h}^{exo}, \hat{c_h}^{exo}, \hat{c_l}^{exo}) + U(\hat{c_h}^{exo}, \hat{c_l}^{exo}, \hat{c_h}^{exo})] + \\
p^2 \cdot [&U({c_h}^{exo}, {c_h}^{exo}, {c_h}^{exo}) + U(\hat{c_h}^{exo}, \hat{c_l}^{exo}, \hat{c_l}^{exo})] \\
U(m_i = 0) = p(1-p) \cdot [&U({c_h}^{exo}, {c_h}^{exo}, {c_l}^{exo}) + U({c_h}^{exo}, {c_l}^{exo}, {c_h}^{exo})] + \\
p^2 \cdot [&U({c_h}^{exo}, {c_h}^{exo}, {c_h}^{exo}) + U(\hat{c_h}^{exo}, \hat{c_l}^{exo}, \hat{c_l}^{exo})]
\end{aligned}
$$

In calculating the difference, we find that $U(\alpha_h | m_i = 1) - U(\alpha_h | m_i = 0) < 0$, therefore a high type would always vote for no monitor.

Combining the two results above, there exists an interval of p where type α_l would choose a monitor whereas type α_h would choose no monitor. Therefore there exists a separating equilibrium with p st. $0 < p < p{*}_{\alpha_l}$ where the above equilibrium holds true. If $p > p{*}_{\alpha_l}$ then there is no separating equilibrium and both types vote for no monitor.