Supporting Information (Online Appendix) for "Channels of Influence or Maps of Behavior?"

Contents

A Community selection procedure

We worked in six communities, a pair of two each from ethnic groups that each represent one of three different histories of traditional leadership in Northern Ghana (Nathan 2019).[1] Communities were chosen non-randomly to both (a) select typical rural villages in Northern Ghana and (b) to examine the most similar possible sets of communities – both within and across ethnic groups – on covariates that we expect affect community social networks separately from traditional leadership institutions. These community-level covariates are measured with geo-referenced enumeration area-level data from the 2010 Ghana census.

Community selection occurred in several stages. We began with a set of 3,588 census enumeration areas covering all of rural Northern Ghana. Given constraints on feasible travel distances, we dropped six extremely remote districts. We also restricted to enumeration areas that were ethnically homogenous and dominated by a single ethnic group ($\geq$ 75% from one group). Most rural villages in Northern Ghana are homogenous (82% of all enumeration areas are above this cutoff). We also restricted the sample by community size (80 to 120 households) and area (15 sq km or less) to focus on rural villages of comparable and modal size.

Within this restricted sample, we use nearest-neighbor Mahalanobis distance matching (with replacement) to select the triplet of communities – one community each with each type of chieftaincy history – that is most similar on a set of covariates. The covariates account for community size, the level of development, major economic activities (e.g., farming vs. trading), ethnic group social structure, and remoteness, all of which could affect community network structures.[2] We first matched the “invented chiefs” communities to the “always chief” communities, defining an indicator of being dominated by an “invented chief” group as the treatment. We then matched the “never recognized” communities to the “always chief” communities, with “never recognized” communities as treated. We then identified all triplets of the two sets of resulting matched pairs that share a common matched “always chief” community and selected the triplet with the smallest maximum Mahalanobis distance between its two component pairs. Next, we selected the remaining 3 communities by finding the best matches within the same ethnic groups for each selected community. This allowed us to hold the selected ethnic groups fixed within chieftaincy type.

Our final selection stage held the spatial compactness of communities fixed, which we assume may also affect the structure of community social networks. Using geo-referenced satellite imagery, we confirmed that all selected communities are spatially compact villages, with all houses in a tight cluster near each other, rather than scattered family homesteads on isolated farms. To ensure balance on this final covariate, we dropped the best match on the other covariates within ethnic groups in two cases and instead select the second-best match, sacrificing marginal balance

---

[1] These are communities with pre-colonially-rooted chieftaincy, communities with chieftaincy imposed in the colonial period, and communities left without formal chieftaincy into the modern period. Variation across histories of chieftaincy is not analyzed in the present paper because of the limited community-level sample size.

[2] They are: population size, number of households, proportion from the community’s majority ethnic group, proportion with access to electricity, proportion with formal or public sector employment, proportion with access to clean drinking water, proportion with english literacy, proportion in homes with formal roofing, proportion in homes owning livestock, whether the ethnic group has patrilineal or duo-lineal inheritance (which may affect family social structures), and the distance (km) to the nearest major town.

on the other covariates to ensure balance on spatial compactness.

B Reasons for dropping the 6th community

Our face-to-face survey interviews were conducted by enumerators entering data via the ODK Collect platform on Android devices. After the first wave (baseline) survey in the second Mamprusi community, one of the enumerators (in a team of 3) mistakenly overwrote all of his completed interview data before being able to reach an area with the mobile data coverage needed to upload it to our server. Because of the continually poor mobile phone connectivity in this remote area, we did not discover this error had occurred until after the experiment had significantly unfolded. It was no longer possible at that point for logistical reasons to re-interview the missing subjects or re-run the experiment. Unfortunately, with one third of households (nodes) now missing, it was also no longer possible to produce accurate social network statistics (e.g., eigenvector centrality) among the remaining respondents. The accumulated donations were still given to the community leader as designed, but this community is dropped in all analyses.

C Demographic characteristics of selected communities

Table OA.1 provides demographic information from the 2010 Ghana census for the set of six selected communities, broken out for the average value for each ethnic group.

Table OA.1: Balance across ethnic groups

Group	Builsa	Konkomba	Mamprusi
Total population	485	785	1090
No. of households	85	85	105
Adult literacy (%)	0.240	0.180	0.086
Electricity (%)	0	0	0
Formal housing (%)	0.246	0.736	0.164
Formal employment (%)	0	0	0
Household engaged in farming (%)	1.000	1.000	0.973
Household owns livestock (%)	1.000	0.807	0.864
Distance to district capital (km)	36.9	58.9	46.2
Patrilineal inheritance (vs. duo-lineal)	0	1	0
N	2	2	2

D Pre-registered experimental analysis

Our pre-analysis plan pre-registered a more complex estimation procedure instead of OLS regressions. Because treated participants could, by design, interact with each other before making donations, the overall effect of treatment is a combination of the direct effect on each treated participant and the indirect effect through possible spillovers from the actions of other treated participants. Despite the random assignment of treatment, the probability treated participants are subject to these spillovers is endogenous to network location: more socially-connected treated participants (i.e., of higher degree) are more likely to have had an alter also assigned to treatment, and thus to interact with other treated participants. Because we are explicitly interested in the effect of allowing for communication among treated participants, our substantive quantity of interest is the combined direct and indirect effect, not the isolated direct effect among respondents not subject to spillovers.

Our pre-analysis plan proposes examining this combined direct and indirect effect by following the approach in Aronow and Samii (2017) for estimating effects in network experiments with interference. We test $H5$ and $H6$ after defining an "exposure mapping" of treatment assigned to treatment received. Participants are classified into three exposure conditions based on their randomly assigned treatment and their existing network location:

- "Direct and indirect treated" $(T_i(1,1))$: units assigned to treatment that also have at least one direct link assigned to treatment
- "Isolated treated" $(T_i(1,0))$: units themselves assigned to treatment, but without direct links assigned to treatment
- "Pure control" $(T_i(0))$: units assigned to control[3]

where $T_i(1,1)$ indicates that participant i was assigned to treatment $T_i(1)$ and has at least one directly linked alter, j, assigned to treatment $T_j(1)$.[4] Our main effect of substantive interest is the comparison of $T_i(1,1)$ vs. $T_i(0)$ – that is, the combined average direct and indirect (spillover) effects of treatment relative to the pure control condition.[5]

We estimate this effect and its interactions with pre-treatment moderators using Weighted Least Squares, weighting by the inverse probability that each participant was assigned to $T_i(1,1)$ and controlling for each participants' network degree to adjust for non-random, pre-treatment differences in participants' probability of being subject to indirect effects via treated alters (Aronow and Samii 2017).[6] All models also control for an indicator for the $T_i(1,0)$ condition and include the same demographic controls and fixed effects as described for the main text.

---

[3] Because control participants could not interact with treated participants, we can collapse $T_i(0,0)$ and $T_i(0,1)$ into a single condition, $T_i(0)$.

[4] After random assignment 45.9% of households were assigned to control ($T_i(0)$), while approximately 5.1% of treatment households were assigned to the $T_i(1,0)$ exposure and 48.8% to the $T_i(1,1)$ exposure.

[5] Other definitions of the exposure conditions would be possible, such as comparisons of $T_i(1,2)$ (treated participants with two treated alters) vs. $T_i(1,1)$ (treated participants with one treated alter). But we forego these more complex analyses due to our limited sample size.

[6] This "exposure probability" of each household is calculated via simulation by re-randomizing treatment assignment 10,000 times and re-calculating each participant's exposure condition based on their position in the observed

Table OA.2: Results for experimental hypotheses

	Outcome variable	*Explanatory variable*	β	p	N
–	Donation (0-5 GHS)	Treatment status T(1,1)	-0.23	0.03^a	391
$H4$	Variance in donation amount (0-5 GHS)	Treatment status T(1,1)	-0.01	0.48	402
$H5a$	Donation (0-5 GHS)	Proximity (SNN) to leader * treatment status T(1,1)	-0.13	0.90	310
		Proximity (MPD) to leader * treatment status T(1,1)	0.03	0.61	311
$H5b$	Same project preference as leader (0,1)	Proximity (SNN) to leader * treatment status T(1,1)	0.03	0.23	326
		Proximity (MPD) to leader * treatment status T(1,1)	0.01	0.73	326
$H6a$	Abs. value of diff. in donation (0-5 GHS)	Proximity (SNN) of dyad * treatment status T(1,1)	-0.006	0.29	15,333
		Proximity (MPD) of dyad * treatment status T(1,1)	0.02	0.29	15,333
$H6b$	Same leader preference in dyad (0,1)	Proximity (SNN) of dyad * treatment status T(1,1)	0.003	0.40	15,346
		Proximity (MPD) of dyad * treatment status T(1,1)	-0.02	0.49	15,346
$H6c$	Same proj. preference in dyad (0,1)	Proximity (SNN) of dyad * treatment status T(1,1)	0.01	0.38	15,154
		Proximity (MPD) of dyad * treatment status T(1,1)	-0.01	0.51	15,154

SNN and MPD refer to "Shared Nearest Neighbors" and "Minimum Path Distance," respectively. P-values calculated via randomization inference, as described in the text. *a*: The first test is for a two-sided p-value vs. a null hypothesis of no effect. All other tests are one-sided, given the directional predictions in the hypotheses.

Because units are not independent, conventional methods for calculating uncertainty will also be biased. As in the main text, we adopt a randomization inference (Fisherian) approach. P-values for the main effect of treatment are calculated via randomization inference for the sharp null hypothesis of no effect for any unit, with the treatment effect estimated via WLS as the test statistic. P-values for the interaction of treatment with pre-treatment moderators (network location) are calculated via randomization inference for a sharp null hypothesis of a constant effect at the overall treatment effect estimated via WLS (without the interaction), and the coefficient on the interaction term as the test statistic. Results for this alternative procedure are in Table OA.2. They are substantively identical to the results in the main text.

We are less interested in the effect of treatment among treated participants without treated alters ($T_i(1,0)$ vs. $T_i(0)$), as it would not test the possible effects of communication, just of waiting to donate. But in Table OA.3 we repeat these analyses for the isolated direct effect of treatment for

---

network. On average, participants had a 50% probability of assignment to $T_i(0)$, a 44% probability of assignment to $T_i(1,1)$, and a 6% probability of assignment to $T_i(1,0)$.

the $T_i(1,0)$ exposure condition. This is the effect of treatment among treated participants who were not linked the network to any other treated participants. The null results for $H5$ and $H6$ remain as before. Importantly, there is no longer a direct effect of treatment on donation amounts, however. This suggests that the negative effect estimated in Table OA.2 and Table 6 emerged through the interaction of treated participants with each other.

Table OA.3: Results for experimental hypotheses with $T_i(1,0)$ as treatment condition

	Outcome variable	*Explanatory variable*	β	p
–	Donation (0-5 GHS)	Treatment status T(1,0)	0.37	0.59^a
$H5a$	Donation (0-5 GHS)	Proximity (SNN) to leader * treatment status T(1,0)	0.47	0.14
		Proximity (MPD) to leader * treatment status T(1,0)	-0.30	0.10
$H5b$	Same project preference as leader (0,1)	Proximity (SNN) to leader * treatment status T(1,0)	-0.02	0.37
		Proximity (MPD) to leader * treatment status T(1,0)	0.13	0.40
$H6a$	Abs. value of diff. in donation (0-5)	Proximity (SNN) of dyad * treatment status T(1,0)	0.20	0.70
		Proximity (MPD) of dyad * treatment status T(1,0)	-0.03	0.55
$H6b$	Same leader preference in dyad (0,1)	Proximity (SNN) of dyad * treatment status T(1,0)	-0.01	0.54
		Proximity (MPD) of dyad * treatment status T(1,0)	-0.01	0.30
$H6c$	Same project preference in dyad (0,1)	Proximity (SNN) of dyad * treatment status T(1,0)	-0.65	1.00
		Proximity (MPD) of dyad * treatment status T(1,0)	0.00	0.52

SNN and MPD refer to “Shared Nearest Neighbors” and “Minimum Path Distance,” respectively. P-values calculated via randomization inference, as described in the text. *a*: The first test is for a two-sided p-value vs. a null hypothesis of no effect. All other tests are one-sided, given the directional predictions in the hypotheses.

E Separate estimates for each type of tie

Tables OA.4 and OA.5 report results for our pre-registered hypotheses[7] when re-running the corresponding specifications for each type of tie (e.g., friendship, relative, etc.), rather than for the union across all four types of ties. Our results are substantively similar for most hypotheses when examining each type of tie separately.

To maintain consistency with the main text, we report estimates for $H5$ and $H6$ using OLS with p-values via randomization inference. These results are also substantively identical for these analyses under the pre-registered estimation procedure (see page SI.4).

Table OA.4: $H1$-$H3$ re-estimated by network domain

	Proximity measure	Full		Help		Politics		Relatives		Friends	
		β	p	β	p	β	p	β	p	β	p
$H1$	–	2.82	<0.001	0.40	0.52	0.67	0.21	2.06	<0.001	0.12	0.85
$H2$	SNN	0.08	0.21	0.08	0.60	-0.23	0.20	0.12	0.15	-0.23	0.14
	MPD	0.07	0.24	-0.02	0.53	0.03	0.28	-0.04	0.34	0.03	0.48
$H3a$	SNN	-0.01	0.80	-0.01	0.75	0.00	0.98	-0.01	0.65	0.00	0.91
	MPD	0.05	0.12	0.02	0.13	0.01	0.32	0.001	0.95	-0.004	0.70
$H3b$	SNN	0.03	<0.01	0.05	<0.001	0.04	0.01	0.03	0.03	0.04	<0.01
	MPD	-0.03	<0.01	-0.01	0.03	-0.01	<0.001	-0.01	0.24	-0.02	<0.001
$H3c$	SNN	0.03	0.02	0.05	<0.01	0.05	0.03	0.05	0.01	0.01	0.43
	MPD	-0.01	0.23	-0.01	0.26	-0.0045	0.17	0.00	0.44	0.00	0.90

SNN and MPD refer to "Shared Nearest Neighbors" and "Minimum Path Distance," respectively. P-values calculated via randomization inference, as described in the text. Hypotheses and specifications mirror that of our main analysis.

Table OA.5: $H5$-$H6$ re-estimated by network domain

	Proximity measure	Full		Help		Politics		Relatives		Friends	
		β	p	β	p	β	p	β	p	β	p
$H5a$	SNN	-0.11	0.85	-0.24	0.80	0.41	0.06	-0.16	0.87	0.36	0.10
	MPD	-0.09	0.24	-0.01	0.44	-0.05	0.12	0.19	0.99	-0.05	0.13
$H5b$	SNN	0.03	0.23	0.03	0.41	-0.04	0.68	0.04	0.19	0.05	0.33
	MPD	0.02	0.71	0.01	0.78	0.00	0.56	0.01	0.62	0.01	0.74
$H6a$	SNN	0.02	0.68	0.09	0.95	0.05	0.72	-0.01	0.43	0.02	0.68
	MPD	-0.05	0.87	-0.02	0.95	0.00	0.38	0.00	0.56	-0.01	0.71
$H6b$	SNN	-0.00	0.51	0.00	0.50	-0.02	0.71	0.02	0.21	0.02	0.15
	MPD	-0.01	0.27	-0.01	0.16	0.00	0.77	-0.01	0.08	-0.02	0.02
$H6c$	SNN	-0.01	0.63	-0.03	0.81	-0.04	0.85	0.02	0.26	-0.01	0.68
	MPD	0.01	0.74	-0.00	0.37	0.01	0.99	-0.00	0.29	-0.01	0.13

SNN and MPD refer to "Shared Nearest Neighbors" and "Minimum Path Distance," respectively. P-values calculated via randomization inference, as described in the text. Hypotheses and specifications mirror that of our main analysis.

---

[7] This does not include $H4$, which does not depend on network measures.

F Main results using the total sum of ties

In our main analyses, we consider two households equally connected if either named the other in one or more of the four "name generator" prompts. In this section, we instead present results using a network weighted by the total number of connections between nodes. The weights are calculated by summing the number of domains in which each edge appears.[8]

To maintain consistency with the main text, we report estimates for Table OA.7 ($H5$, $H6$) using OLS with p-values via randomization inference. These results are also substantively identical for these analyses using the pre-registered estimation procedure instead (see page SI.4).

Table OA.6: Results for non-experimental hypotheses *with networks weighted by edge frequency*

	Outcome variable	*Explanatory variable*	β	p	95% CI
$H1$	Donation (0-5 GHS)	Betweenness Centrality	0.001	<0.01	(0.000, 0.002)
		Degree Centrality	0.07	<0.01	(0.035, 0.096)
$H2$	Donation (0-5 GHS)	Proximity (SNN) to leader	0.006	0.98	(-0.58, 0.59)
		Proximity (MPD) to leader	0.04	0.43	(-0.06, 0.14)
$H3a$	Abs. value of diff. in donations (0-5 GHS)	Proximity (SNN) of dyad	-0.01	0.80	(-0.05, 0.04)
		Proximity (MPD) of dyad	-0.001	0.96	(-0.05, 0.05)
$H3b$	Same leader preference in dyad (0,1)	Proximity (SNN) of dyad	0.03	<0.01	(0.01, 0.05)
		Proximity (MPD) of dyad	-0.05	<0.001	(-0.08, 0.02)
$H3c$	Same proj. preference in dyad (0,1)	Proximity (SNN) of dyad	0.03	<0.10	0.003, 0.046)
		Proximity (MPD) of dyad	-0.01	0.38	(-0.04, 0.02)

---

[8] Note that we use both degree and betweenness centrality (where applicable) instead of eigenvector centrality because the *tnet* package in R used to work with weighted networks does not have a method for calculating eigenvector centrality. Pg. SI.16 shows that these alternative measures of centrality are highly correlated and effectively interchangeable with eigenvector centrality in our application.

Table OA.7: Results for experimental hypotheses *with networks weighted by edge frequency*

	Outcome variable	*Explanatory variable*	β	p
$H5a$	Donation (0-5 GHS)	Proximity (SNN) to leader * treatment status (0,1)	-0.10	0.83
		Proximity (MPD) to leader * treatment status (0,1)	-0.19	0.04
$H5b$	Same project pref. as leader (0,1)	Proximity (SNN) to leader * treatment status (0,1)	0.03	0.23
		Proximity (MPD) to leader * treatment status (0,1)	0.01	0.62
$H6a$	Abs. value of diff. in donation (0-5 GHS)	Proximity (SNN) of dyad * treatment status (0,1)	0.02	0.70
		Proximity (MPD) of dyad * treatment status (0,1)	-0.06	0.93
$H6b$	Same leader pref. in dyad (0,1)	Proximity (SNN) of dyad * treatment status (0,1)	0.00	0.47
		Proximity (MPD) of dyad * treatment status (0,1)	-0.01	0.16
$H6c$	Same proj. pref. in dyad (0,1)	Proximity (SNN) of dyad * treatment status (0,1)	-0.01	0.68
		Proximity (MPD) of dyad * treatment status (0,1)	0.01	0.67

SNN and MPD refer to “Shared Nearest Neighbors” and “Minimum Path Distance,” respectively. P-values calculated via randomization inference, as described in the text. *a*: The first test is for a two-sided p-value vs. a null hypothesis of no effect. All other tests are one-sided, given the directional predictions in the hypotheses.

G Incorporating unmatched alters

We also present alternative estimates using alternative measures of the social networks that incorporate information lost by dropping alters that were not matched back to an interviewed household. In the "closed" networks commonly used in the literature, information about unmatched, uninterviewed alters is simply dropped; all nodes in the network are interviewed respondents. But when two households name the same uninterviewed alter, this indicates the presence of an indirect, second-order connection between those households traveling through that unmatched alter that is being ignored in our main specifications. In our alternative specification, we include second-order ties between households via unmatched alters by using a weighted network. Direct ties between households have double the weight of indirect ties via unmatched alters.

Table OA.8 and Table OA.9 present our main results with the new measures of centrality and distance from the weighted network incorporating indirect ties.[9] Our results are generally robust to including these unmatched ties.

To maintain consistency with the main text, we report estimates for Table OA.9 ($H5$, $H6$) using OLS with p-values via randomization inference. These results are also substantively identical for these analyses when instead using the pre-registered estimation procedure (see page SI.4).

Table OA.8: Results for non-experimental hypotheses *including unmatched alters*

	Outcome variable	*Explanatory variable*	β	p	95% CI
$H1$	Donation (0-5 GHS)	Eigenvector centrality	0.05	<0.001	(0.02, 0.08)
$H2$	Donation (0-5 GHS)	Proximity (SNN) to leader	0.09	0.09	(-0.02, 0.20)
		Proximity (MPD) to leader	0.004	0.95	(-0.13, 0.14)
$H3a$	Abs. value of diff. in donations (0-5 GHS)	Proximity (SNN) of dyad	-0.01	0.70	(-0.05, 0.04)
		Proximity (MPD) of dyad	0.01	0.77	(-0.04, 0.06)
$H3b$	Same leader preference in dyad (0,1)	Proximity (SNN) of dyad	0.03	<0.01	(0.01, 0.05)
		Proximity (MPD) of dyad	-0.07	<0.001	(-0.10, -0.04)
$H3c$	Same project preference in dyad (0,1)	Proximity (SNN) of dyad	0.03	0.02	(0.00, 0.05)
		Proximity (MPD) of dyad	-0.02	0.13	(-0.05, 0.01)

SNN and MPD refer to "Shared Nearest Neighbors" and "Minimum Path Distance," respectively. Estimates from OLS regressions, as described in the text. P-values and confidence intervals calculated from standard errors adjusted for dyadic data for $H3a$-$H3b$, following Aronow et al. (2015).

---

[9] Whereas $H1$ is tested using eigenvector centrality in the paper, available packages in R do not calculate eigenvector centrality over weighted networks. This forces us to use degree centrality instead in these estimates.

Table OA.9: Results for experimental hypotheses *including unmatched alters*

	Outcome variable	*Explanatory variable*	β	p
–	Donation (0-5 GHS)	Treatment status (0,1)	-0.24	0.02^a
$H5a$	Donation (0-5 GHS)	Proximity (SNN) to leader * treatment status (0,1)	-0.11	0.84
		Proximity (MPD) to leader * treatment status (0,1)	0.02	0.54
$H5b$	Same project pref. as leader (0,1)	Proximity (SNN) to leader * treatment status (0,1)	0.03	0.24
		Proximity (MPD) to leader * treatment status (0,1)	0.03	0.74
$H6a$	Abs. value of diff. in donation (0-5 GHS)	Proximity (SNN) of dyad * treatment status (0,1)	0.02	0.68
		Proximity (MPD) of dyad * treatment status (0,1)	-0.02	0.64
$H6b$	Same leader pref. in dyad (0,1)	Proximity (SNN) of dyad * treatment status (0,1)	-0.00	0.52
		Proximity (MPD) of dyad * treatment status (0,1)	0.00	0.57
$H6c$	Same proj. pref. in dyad (0,1)	Proximity (SNN) of dyad * treatment status (0,1)	-0.01	0.70
		Proximity (MPD) of dyad * treatment status (0,1)	0.03	0.88

SNN and MPD refer to "Shared Nearest Neighbors" and "Minimum Path Distance," respectively. P-values calculated via randomization inference, as described in the text. *a*: The first test is for a two-sided p-value vs. a null hypothesis of no effect. All other tests are one-sided, given the directional predictions in the hypotheses.

H Balance by treatment condition

All of our OLS specifications in the main text include a set of household-level controls found in the left-hand column of Table OA.10. Table OA.10 presents the p-value for the difference-in-means between the treatment and control groups for each of these covariates.

Table OA.10: Covariate balance by treatment status

Group	Donate immediately (Control)	Donate w/ delay (Treatment)	p
Eigenvector centrality	0.08	0.08	0.90
HH head age	43.88	43.53	0.82
Number of adults in HH	3.33	3.45	0.53
Number of children in HH	4.35	4.40	0.88
HH head attended Junior Sec School (%) or above	0.10	0.15	0.12
HH head regularly travels (%) from community	33.90	27.95	0.19
Assets index (out of 9)	3.11	3.20	0.53
HH head employed in skilled trade (%)	0.03	0.05	0.26
HH head gender (% Female)	0.20	0.19	0.83
N	202	236	

I Assets index

The assets index used as a covariate throughout the paper consists of the sum of nine common items. Respondents were asked whether they had (1) livestock, (2) electricity, (3) a radio, (4) a bicycle, (5) a motorcycle, (6) a gas stove, (7) a "yam phone" (basic mobile phone), and (8) a smartphone. Last, enumerators observed whether a respondent's (9) home walls were made of concrete (rather than mud). Figure OA.2 shows the distribution of the this index in the survey sample.

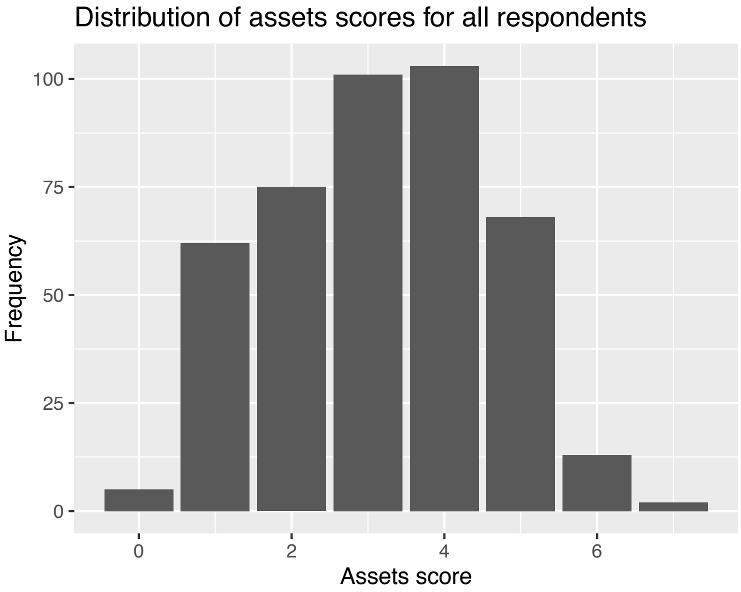

Figure OA.2: *Histogram of assets index for all respondents (households).*

J Main results without controls

Tables OA.11 and OA.12 report the main results without controls for $H1$-$H3c$ and $H5a$-$H6c$ respectively. The results are generally consistent with our main specification. The most notable deviation in Table OA.11 is for $H3c$ when using the "shared nearest neighbors" measure, which is now no longer significant at conventional levels. In Table OA.12 we show that all null results stand. These controls, which we show are balanced across treatment and control groups in Table OA.10, allow us to reduce the standard errors and improve the precision of the estimates reported in the paper.

To maintain consistency with the main text, we report estimates for Table OA.12 ($H5$, $H6$) using OLS with p-values via randomization inference. These results are also substantively identical for these analyses using the pre-registered estimation procedure instead (see page SI.4).

Table OA.11: Results for non-experimental hypotheses *without controls*

	Outcome variable	*Explanatory variable*	β	p	95% CI
$H1$	Donation (0-5 GHS)	Eigenvector centrality	2.98	<0.001	(1.48, 4.48)
$H2$	Donation (0-5 GHS)	Proximity (SNN) to leader	0.09	0.14	(-0.03, 0.22)
		Proximity (MPD) to leader	0.06	0.33	(-0.06, 0.19)
$H3a$	Abs. value of diff. in donations (0-5 GHS)	Proximity (SNN) of dyad	0.00	0.91	(-0.05, 0.04)
		Proximity (MPD) of dyad	0.05	0.12	(-0.01, 0.12)
$H3b$	Same leader preference in dyad (0,1)	Proximity (SNN) of dyad	0.02	0.02	(0.00, 0.04)
		Proximity (MPD) of dyad	-0.03	0.02	(-0.05, 0.00)
$H3c$	Same project preference in dyad (0,1)	Proximity (SNN) of dyad	0.02	0.17	(-0.01, 0.04)
		Proximity (MPD) of dyad	-0.01	0.38	(-0.03, 0.01)

SNN and MPD refer to "Shared Nearest Neighbors" and "Minimum Path Distance," respectively. Estimates from OLS regressions, as described in the text. P-values and confidence intervals calculated from standard errors adjusted for dyadic data for $H3a$-$H3b$, following Aronow et al. (2015).

Table OA.12: Results for experimental hypotheses *without controls*

	Outcome variable	*Explanatory variable*	β	p
–	Donation (0-5 GHS)	Treatment status (0,1)	-0.20	0.05^a
$H5a$	Donation (0-5 GHS)	Proximity (SNN) to leader * treatment status (0,1)	-0.12	0.87
		Proximity (MPD) to leader * treatment status (0,1)	-0.10	0.23
$H5b$	Same project pref. as leader (0,1)	Proximity (SNN) to leader * treatment status (0,1)	0.03	0.26
		Proximity (MPD) to leader * treatment status (0,1)	0.04	0.83
$H6a$	Abs. value of diff. in donation (0-5 GHS)	Proximity (SNN) of dyad * treatment status (0,1)	0.01	0.64
		Proximity (MPD) of dyad * treatment status (0,1)	-0.05	0.84
$H6b$	Same leader pref. in dyad (0,1)	Proximity (SNN) of dyad * treatment status (0,1)	-0.00	0.52
		Proximity (MPD) of dyad * treatment status (0,1)	-0.01	0.27
$H6c$	Same proj. pref. in dyad (0,1)	Proximity (SNN) of dyad * treatment status (0,1)	-0.00	0.61
		Proximity (MPD) of dyad * treatment status (0,1)	0.01	0.76

SNN and MPD refer to "Shared Nearest Neighbors" and "Minimum Path Distance," respectively. P-values calculated via randomization inference, as described in the text. *a*: The first test is for a two-sided p-value vs. a null hypothesis of no effect. All other tests are one-sided, given the directional predictions in the hypotheses.

K Robustness to alternative centrality measures

Table OA.13 presents alternative specifications using either degree or betweenness centrality as an alternative measure to eigenvector centrality. Our Pre-Analysis Plan specified eigenvector centrality as our preferred measure, so that is why we prioritize the eigenvector centrality results in the main text. The results for $H1$ and $H2$ are unchanged, with the exception of the estimate for $H2$ using the shared nearest neighbors measure of proximity and betweenness centrality.

Table OA.13: Results for non-experimental hypotheses *with alternative centrality*

	Outcome variable	*Explanatory variable*	*Centrality measure*	β	p	95% CI
$H1$	Donation (0-5 GHS)	Betweenness Centrality	Betweenness	0.00	<0.001	(0.00, 0.00)
		Degree Centrality	Degree	0.03	<0.001	(0.02, 0.05)
$H2$	Donation (0-5 GHS)	Proximity (SNN) to leader	Betweenness	0.13	0.01	(0.03, 0.23)
		Proximity (MPD) to leader	Betweenness	-0.00	0.99	(-0.11, 0.11)
		Proximity (SNN) to leader	Degree	0.08	0.16	(-0.03, 0.19)
		Proximity (MPD) to leader	Degree	0.05	0.42	(-0.07, 0.16)

L Interview order in the control group

In the main text we assume that influence and network spillovers should not occur in the control group as respondents did not have time to coordinate on behavior in the game between being first introduced to the activity and contributing, both of which occurred during the same survey interview. However, it remains possible that "chatter" could grow starting with the completion of the first interview in each village, such that control group respondents might be subject to or exercise some form of influence prior to being interviewed.

To test for this, we examine minutes between the completion of the first interview and the initiation of every other control group interview in each village. Table OA.14 lists OLS models in which the outcomes are donation amount (column 1) and expected donations of others (column 2) and the main predictor is minutes between when the first interview in each village began to when each control participant's interview began. We include the same set of controls as Table 5 of the main text. We find no differences in donations or beliefs about others' donations among control respondents who were interviewed later within their communities, inconsistent with interference passing among control respondents.

Table OA.14: Control group behavior by time from first interview

	1	2
Outcome:	Donation amount (GHS)	Expected donation of others (GHS)
Minutes from first interview in community	0.0002 (0.0004)	0.0005 (0.0005)
Community fixed effects	Y	Y
Demographic controls	Y	Y
N	195	127
adj. R^2	0.0971	0.1224

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Subset to control groups respondents only. Column 2 is NA if respondents reported they did not know others' likely donation amounts. All models are OLS with standard errors in parentheses.

M Accuracy of expectations about others' behavior, control group

Table OA.15: Accuracy of expected peer behavior, control participants only

Community	Avg error from median donation (GHS)	"Don't know" others' donations	Correctly identify modal project	"Don't know" modal project	Correctly idenitify modal leader	"Don't know" modal leader	N (control)
Builsa 1	-0.50	52.6%	60.5%	0.0%	55.0%	47.4%	38
Builsa 2	0.23	59.4%	60.0%	6.25%	57.1%	12.5%	32
Konkomba 1	0.19	8.5%	57.5%	0.0%	71.4%	10.6%	47
Konkomba 2	0.00	30.6%	52.8%	0.0%	69.6 %	36.1%	36
Mamprusi 1	-0.75	33.3%	82.5%	4.8%	100%	11.9%	42
Overall	-0.15	34.9%	62.8%	2.1%	73.3%	23.1%	195

Note: values in columns 1, 3, and 5 are among the participants who did not answer "don't know."

While some control participants expressed uncertainty about how others would donate (34.9%), most held fairly accurate expectations about their peers' behavior in the activity. This suggests that participants seeking to apply a rule of conditional cooperation could act from a position in which they were already reasonably able to infer other community members' likely behavior, even absent any communication or interaction.

Table OA.15 lists the accuracy of control group participants' priors about the behavior of their peers for each of the three decisions to be made: donation amount, choice of project type (e.g., water vs. school), and choice of leader to manage the project (e.g., chief vs. assemblyman). There is clear variation in the accuracy of priors across the five communities, with the greatest uncertainty about others' behavior in the Builsa communities, which are notably the same two communities with the greatest number of socially-disconnected households in Figure 1 in the main text. But, on average, two-thirds of participants felt they could guess how much others would donate, and those who guessed only missed the correct median donation by an average of 0.15 GHS. Similarly large majorities accurately guessed what project and which leader would be chosen.

N Conditional cooperation and centrality in the control group

The OLS model in Table OA.16 regress control respondents' donation amounts on the interaction of their expectation of others' donation amounts and their eigenvector centrality, including the same individual-level controls and fixed effects as the other analyses. The interaction is significant at the $p < 0.01$ level.

Table OA.16: Conditional cooperation and centrality in the control group

Outcome:	Donation amount (GHS)
Expected donation of others (GHS)	0.07
	(0.13)
Eigenvector centrality	-3.69
	(2.68)
Expected donation of others (GHS)	3.24^{**}
* eigenvector centrality	(1.23)
Community fixed effects	Y
Demographic controls	Y
N	127
adj. R^2	0.25

Standard errors in parentheses

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Subset to control participants only.

O H4 separately by community

In Table OA.17 we show estimates of $H4$ at the community-level. We apply the same estimation method described in the main text to subsets of the data (divided by community). Only in the second Konkomba community is the change in variance signed in the same direction as predicted by $H4$ and significant at the $p = 0.1$ level in a one-sided test.

Table OA.17: Results for H4 by community

Case	β	p
$H4$ overall	-0.01	0.48
Builsa 1	-0.44	0.16
Builsa 2	0.12	0.47
Konkomba 1	0.06	0.44
Konkomba 2	-0.70	0.09
Mamprusi 1	-0.45	0.20

P-values calculated via randomization inference, as described for $H4$ in the text.

P Additional tests for a wealth effect

We further rule out that our main result for network centrality ($H1$) can be explained by more central individuals being wealthier using an additional, stricter assets measure than the nine-item index described in on page SI.13. In this alternative measure, we include only the three highest-value assets measured on the survey: having a smartphone, motorcycle, and electricity. Respondents with these three particular assets should be the wealthiest in the survey sample. Table OA.18 reports our test for $H1$ using the same estimation procedure described in the text, but with this alternative assets measure instead of the regular control. The result remains the same, suggesting wealth is unlikely to account for the relationship between network centrality and cooperation.

Table OA.18: Results for H1 *with alternative assets index*

	Outcome variable	*Explanatory variable*	β	p	95% CI	N
$H1$	Donation (0-5 GHS)	Eigenvector centrality	2.88	<0.001	(1.32, 4.43)	402

Q Differential benefits to more central players?

Another possibility for why more central participants already donate more in the control group, absent social influence, is that they rationally expect to be more likely to benefit from a local public good. We believe this alternative is unlikely for three reasons.

First, the main projects at stake mostly have locally non-excludable benefits that (if properly funded) should have been expected to benefit all participants similarly, especially once already controlling for participants' wealth and other indicators of socio-economic status (as we do in all analyses in the text).[10] Table OA.19 lists the distribution of donors' preferred projects. The main projects at stake in the public goods game were: electricity, especially raising funds for the purchase of poles to enable the community to be connected to the grid; drinking water access, especially raising funds for repairs to the community's public borehole; and funding for the local health clinic. None of these are likely to have highly differential benefits across users within these small, rural communities.

Second, our null result for $H2$ is inconsistent with this alternative. One of the most plausible ways in which ostensibly locally non-excludable goods like those in Table OA.19 could still have differentially greater benefits for some community members than others is if the leaders tasked with creating the good restrict access to a subset of community members, or divert the accumulated donations to favored individuals within the community. If either were true, we would expect participants who are closer in the network to the leader expected to manage the donations to expect to be less likely to be excluded from the benefits. Under the alternative explanation, these individuals should then be more likely to donate in anticipation of greater benefits. But for $H2$, we find that participants closer to these leaders in the network donated no more than those further away.

Third, in a purely rationalist framework, Olson (1965) argues that individuals who stand to benefit differentially more than others from a public good will forego free-riding (cooperate) if their private contribution can have a large enough marginal effect on the outcome to provide the good on their own. But individual donations in our study were capped at 5GHS (10GHS with matching) – \$1.25 (\$2.50) – far too small on the margins to affect whether any good would actually be

---

[10] Moreover, more central players in the networks are marginally wealthier in terms of assets. To the extent that there is diminishing marginal utility of wealth, it is instead more plausible that the most network central players are those who stood to benefit *least* on average from local public goods that would provide a free service (e.g., clean water) that otherwise must be purchased privately.

Table OA.19: Preferences by community

Good	Builsa 1	Mamprusi 1	Konkomba 1	Konkomba 2	Builsa 2
Agriculture	1	1	2	2	2
Education	0	3	1	12	3
Electricity	58	1	59	0	13
Health	0	1	9	43	43
Infrastructure	0	6	3	1	5
Religious	0	0	1	4	0
Water	22	58	21	12	2
Other/NA	7	21	1	8	5

Note: Values express frequency of response after coding from an open-form response.

delivered. Absent consideration of cooperative norms or players' underlying types (our preferred explanations), players who stood to benefit differentially more from the good should still have rationally free-rided.

R Did treated participants simply consume the endowment?

Some treated participants may have needed to spend the endowment on a more pressing need that came up during the waiting period. Indeed, our covariate measuring household assets positively predicts donation amounts, consistent with poorer participants keeping more of the endowment.[11] But this is unlikely to account for our results for several reasons.

First, the assets index does not significantly interact with treatment, with poorer participants in the treated group no more likely to keep the money than similarly poor participants in the control group. Moreover, only 5% (10) of participants in the treatment group made a 0 GHS donation, inconsistent with many treated participants consuming their full endowments.

Second, we cast doubt on this alternative explanation by examining variation in the length of the wait between interviews for the treated participants. On average, the follow-up interview occurred roughly one and a half days (33.9 hours) after the initial interview, but this varied for logistical reasons from a few participants who were (mistakenly) reinterviewed later the same day[12] to a few only re-contacted 3 days later (up to 67 hours later). If treated participants donated less overall because they became tempted to consume the endowment in the interim, donation amounts should be systematically lower among participants who had to wait longer and had more opportunities for other pressing expenses to come up. But we find no relationship between hours waited and amount donated in Table OA.20. We regress hours between the conclusion of the first interview and the

---

[11] The assets index includes owning a radio, livestock, an electricity connection, a bicycle, a motorbike, an improved stove, basic mobile phone, and smartphone.

[12] Seven treated participants (3%) were incorrectly re-interviewed later on the same day as the initial interview.

start of the follow-up interview on the amount donated, while including the full battery of controls used throughout the main text.

Table OA.20: Impact of wait-time on treatment group behavior

	1
Outcome:	Donation amount (GHS)
Wait hours	-0.001
	(0.005)
Community fixed effects	Y
Demographic controls	Y
N	207
adj. R^2	0.334

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Subset to control group respondents only. OLS with standard errors in parentheses.

S Chiefs' contacts in the treated group

In Table OA.21 we regress whether a respondent in the treatment group spoke to the chief on two measures of network distance from the chief. We include the full set of controls used in our main specifications. At least for the path distance measure, chiefs were more likely to talk about the activity with their closer social alters.

Table OA.21: Impact of network distance from chief on probability of speaking with chief

Outcome:	Spoke to Chief (0,1)	Spoke to Chief (0,1)
Path Distance from the plurality chosen leader	$-0.55^{\dagger}$	
	(0.30)	
Shared nearest neighbors with plurality chosen leader		0.20
		(0.27)
Community fixed effects	Y	Y
Demographic controls	Y	Y
N	162	161

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Subset to treatment participants only. Logistic regressions with standard errors in parentheses.

T Conditional cooperation in the treatment group

Table OA.22 presents the results of regressing treatment respondents' own donation behavior on their expectations of others' donation amounts. On average, for each additional cedi that a treatment respondent expected others to donate, she donated an additional 0.22 cedis to the public good ($p < 0.01$).

Table OA.22: Conditional Cooperation in the Treatment Group

Outcome:	Donation amount (GHS)
Expected donation of others (GHS)	0.22^{**}
	(0.08)
Community fixed effects	Y
Demographic controls	Y
N	123
adj. R^2	0.46

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Subset to treatment participants only, dropping those who report they don't know how much others will donate (39%). OLS with standard errors in parentheses.

U Interaction of treatment and eigenvector centrality

The relationship between eigenvector centrality and donation amount does not vary with treatment. More central participants donated similarly more in both the control and treatment groups. Table OA.23 provides an OLS regression table for this interaction. The p-value on the interaction term calculated via randomization inference is $p = 0.62$.

Table OA.23: Interaction between treatment and eigenvector centrality

Outcome:	Donation amount (GHS)
Treatment (0,1)	-0.32^{*}
	(0.16)
Eigenvector centrality	$2.23^{\dagger}$
	(1.15)
Treatment * Eigenvector	1.08
	(1.51)
Community fixed effects	Y
Demographic controls	Y
N	402
adj. R^2	0.19

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Subset to treatment participants only, dropping those who report they don't know how much others will donate (39%). OLS with standard errors in parentheses.

V Treatment effects on expectations of others' behavior

In Table OA.24 we estimate the effects of treatment on three additional outcome variables that indicate the possible effects of communication. The first outcome is the accuracy of participants' final guess of the amount others would donate, measured as the absolute value of the difference between each participants' guess and the median donation in their community. The second outcome is whether participants instead indicated that they "didn't know" what others would donate in their final statement of expectations of others' behavior. The third outcome is how much participants expected others to donate (with "don't know" set to NA). In each case, these are expectations measured right before participants made their donation: at the end of the first wave interview for control participants and end of the second wave interview for treated participants.

To maintain consistency with the main text, these estimates are from OLS regressions with the same controls and p-values calculated via randomization inference. These are 2-sided tests against null hypotheses of no effect for any unit. An alternative set of estimates using the pre-registered approach from Aronow and Samii (2017) is substantively same except for the "don't know" outcome. Using the pre-registered approach instead, we find that treated participants were 7.1 percentage points more likely to express uncertainty about how others would behave (albeit only at the $p < 0.1$ level). This estimate is not statistically significant using the alternative OLS specification presented here.

Table OA.24: Effects of treatment on expectations of others' behavior

Outcome:	β	p
Accuracy of expected donation of others	-0.10	0.25
"Don't know" others' donation amount	0.04	0.44
Amount others expected to donate	-0.10	0.33