It Takes a Village: Peer Effects and Externalities in Technology Adoption

Supplementary Information

September 29, 2019

Contents

List of Tables

List of Figures

1 Glossary of network concepts

A network, often called a *graph*, is a collection of nodes and of ties between these nodes. We write the graph $g = (G, N)$, where N is the set of nodes, and G is the set of ties, and a tie is a pair $(i,j), i,j \in N$. Networks can also be represented by an $(N \times N)$ *adjacency matrix* m, where $m_{ij} = 1$ if there is a tie from i to j, and $m_{ij} = 0$ otherwise. The *size* of g is its amount of nodes.

A graph can be *directed* or *undirected*. In the former case, there is a distinction between a tie from i to j and a tie from j to i. That is, we do not require that m is symmetric. In the latter case, there is no distinction, and we require that m is symmetric. In what follows, we define the network concepts used in the paper in the case of an undirected network.

- *Neighbor*: j is a neighbor of i if they are connected; that is, if $(i,j) \in G$. The neighborhood of i is the set of i's neighbors.
- *Degree*: the degree d_i of i is the number of neighbors i has. That is, $d_i = \sum_{j \neq i} m_{ij}$.
- *Isolate*: i is an isolate if it has a degree of 0.
- *Density*: captures the amount of ties in g, relative to its size. A network of size n has $T_g = n(n-1)/2$ ties. Let $t_g = \sum_{i<j} m_{ij}$ be the amount of ties in g. The density of g is $D_g = t_g/T_g$.
- *Clustering coefficient*: the extent to which the friends of i are friends with each other. Formally, it is the amount of triangles in i's neighborhood normalized by the amount of triangles in i's neighborhood. It writes $c_i = \sum_j \sum_k m_{ij} m_{ik} m_{jk} / \sum_j \sum_k m_{ij} m_{ik}$, with $i \neq j, i \neq k, j < k$.
- *Path*: a path between i and j is a route from i to j on the graph g. Formally, it is a sequence of ties $(i_1, i_2), (i_2, i_3), ..., (i_{K-1}, i_K)$ such that $(i_k, i_{k+1}) \in G$ for each $k \in \{1, ..., K-1\}$, with $i_1 = i, i_K = j$, and each node in the sequence $i_1, ..., i_K$ is distinct.
- *Connected* graph: a graph is connected if there is a path between any $i,j \in N$
- *Path length*: the number of steps it takes to get from i to j on some path. Formally, the length of path $p = (i_1, i_2), (i_2, i_3), ..., (i_{K-1}, i_K)$ is $K-1$.
- *Distance*: the distance l_{ij} between i and j is the length of the shortest path between i and j.
- *Closeness centrality*: how close is node i from the rest of the graph? The closeness centrality of i is the mean distance between i and all other nodes of the graphs. It writes $L_i = \sum_{j \neq i} l_{ij}/(N-1)$. The concept is not well-defined when the graph g is not connected.
- *Betweenness centrality*: how much do people have to go through node i? Betweenness centrality is, for any $j,k \neq i$, the amount of shortest paths that go through i. The concept is not well-defined when the graph g is not connected.

2 Additional information on the setting

This section provides additional information relative to section 2 in the paper. Specifically, we report the regression we used to select the villages included in the study, discuss how Figure 2 was constructed, how individual-level covariates were measured, and how network ties were constructed (e.g., survey question verbatim; how we deal with missingness). Finally, we provide additional information on the distribution of ties across network types.

2.1 Village selection

Table 1 shows the results of a set of OLS regressions where measures of adoption are regressed on village-level factors. Figure 1 shows plots of the residuals. For budget reasons, we decided to survey 16 villages. We selected the 8 villages with the lowest and highest residuals that were accessible to our survey teams. Due to a replacement that took place during fieldwork, the sample then included one more high-uptake village and one less low-uptake village.

	Dependent variable: Number of messages per 100 inhabitants
GAPP community meeting	3.150
	(4.264)
GAPP dialogue meeting	−6.951
	(4.367)
adult population	-4.092^{**}
	(1.648)
pct. secondary education	3.001
	(3.217)
pct. non-agriculture	0.925
	(2.046)
trading center	−2.055
	(3.063)
ethnic fractionalization	−1.610
	(1.361)
distance to health center	−3.075
	(2.097)
distance to school	−4.132
	(2.818)
distance to Arua	2.929
	(2.566)
Constant	17.353^{***}
	(3.724)
Observations	86
R^2	0.157
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 1: **Village selection.** Regression used to select high and low-uptake villages. Standard errors are clustered by health center (i.e. all villages using the same health-center are grouped into a single cluster). The variables GAPP community meeting and GAPP dialogue meeting refer to villages where initial meetings introducing the program and follow-up meetings were held, respectively. The variable trading center is a binary variable that takes a value of 1 if the village has a trading center.

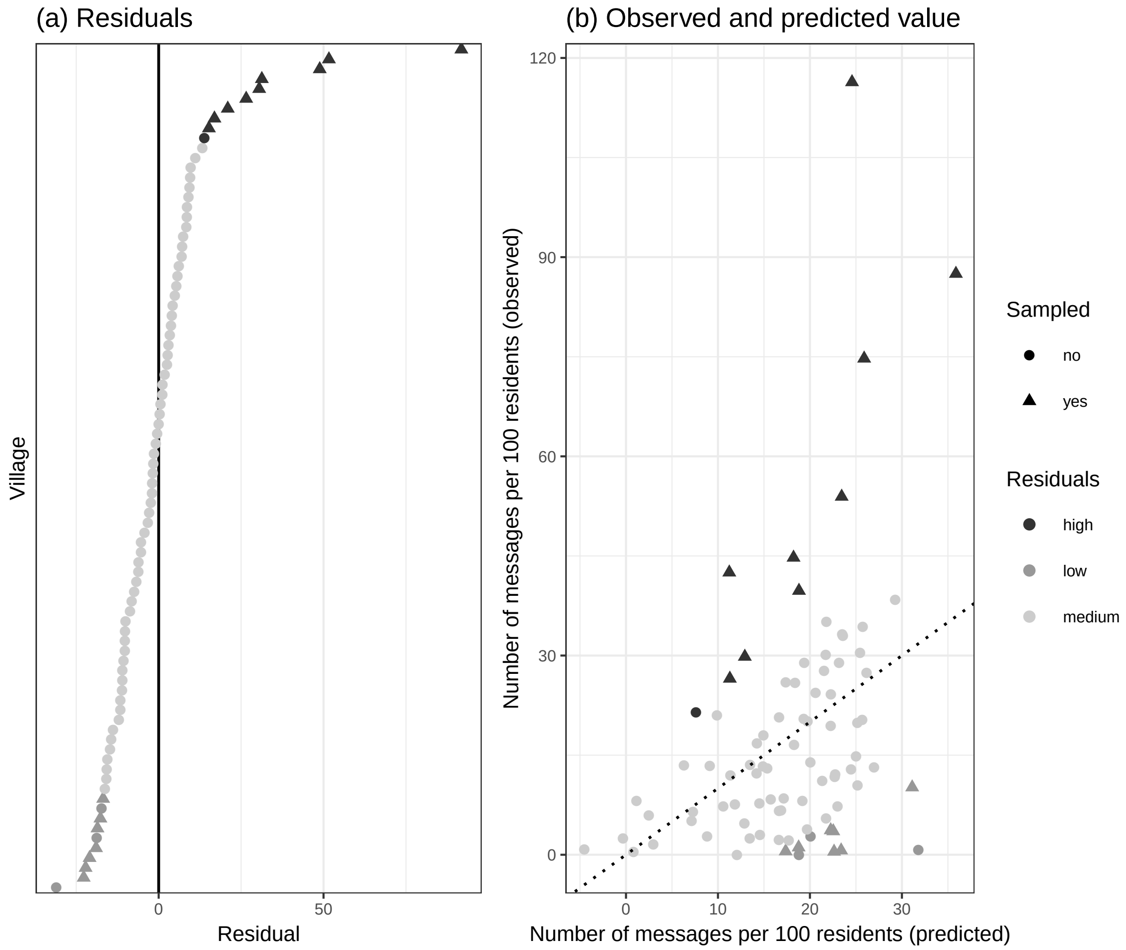

Figure 1: **Residuals of village selection regression.** Colors report the 10 villages with highest and lowest residuals. Triangles report, out of these 20 villages, the ones that were included in the sample.

2.2 Construction of Figure 2

The right hand side of Figure 2 groups villages into high- and low-uptake clusters using the grouping originating from our village-selection regression reported in Table 1 above. Using this grouping, we simply count the number of relevant messages sent by each of these groups for each month and normalize by the size of the population. We then plot the loess fit of each of these two time-series.

We construct the left hand side of Figure 2 in the exact same way, with the exception that we cluster these villages endogenously using a gaussian mixture model with two mixture components estimated using the Expectation Maximization (EM) algorithm. Let $\pi_j \in \{0,1\}$ be the latent type of village j, and let y_{jt} be the number of relevant messages per 100 adult residents received from village j during month t. Our specification is

$$y_{jt} = 1\{\pi_j = k\} \left(\beta_{0k} + \beta_{1k} t + \beta_{2k} t^2 + \epsilon_{kjt}\right),$$

with $\epsilon_{kjt} \sim N(0, \sigma_k^2)$, a cluster-specific error term. We picked a quadratic specification because it would allow for the inverse U-shaped pattern we find for high-uptake villages in the right hand side panel. The output of such mixture model is (1) a series of 8 parameters $\theta = \{\beta_{00}, ..., \beta_{40}, \sigma_0, \beta_{01}, ..., \beta_{41}, \sigma_1\}$, and (2) a series of mixture weights; that is, for any village j, a weight $w_j \Pr(\pi_j = 1|y_j, \theta)$. We assign village j to cluster 1 whenever $w_j \geq .5$, and assign it to cluster 0 otherwise. We then label as high-uptake the cluster that has the highest number of messages per 100 adult residents and as low-uptake the other cluster. We finally construt the left hand side of Figure 2 as its right hand side.

2.3 Description of Covariates

We begin by describing the measurement individual-level variables that are used as controls in (some of) our regression models. These covariates include: *age*; a *female* indicator; *secondary education*, a binary variable that equals 1 if the respondent attained at least secondary education; and *income*, a subjective wealth measure ranging from 1 (low) to 5 (high). The variable *use phone* is a binary variable that equals 1 if the respondent has used a mobile phone in the past 12 months. *Leader* is a binary variable that equals 1 if the respondent occupies a formal leadership position within the village. *Political participation* is a summary index aggregating across recent political actions. We consider attending a village meeting, contributing money to a village project or a village member, contributing labor to a village project, reporting a problem to a village leader, and reporting a problem to the local government, in the past 12 months. The summary index is constructed following the method proposed by Anderson (2008), which gives more weight to more separating components of the index. *Pro-sociality* is a behavioral proxy-measure of care for the community; it is measured as the amount contributed in a standard dictator game. Finally, *attend meeting* indicates whether the respondent attended the GAPP's community meetings, in which the U-Bridge platform has been introduced

2.4 Network Construction

First, we provide verbatim excerpt from our in-person survey used to construct adjacency matrices capturing within-village network ties.

"In each of the following questions, we will ask you to think about people in your community and their relationships to you."

- **Family**: "Think about up to five family members in this village not living in your household with whom you most frequently spend time. For instance, you might visit one another, eat meals together, or attend events together."
- **Friends**: "Think about up to five of your best friends in this village. By friends I mean someone who will help you when you have a problem or who spends much of his or her free time with you. If there are less than five, that is okay too."
- **Lender**: "Think about up to five people in this village that you would ask to borrow a significant amount of money if you had a personal emergency."
- **Problem solver**: "Imagine there is a problem with public services in this village. For example, you might imagine that a teacher has not come to school for several days or that a borehole in your village needs to be repaired. Think about up to five people in this village whom you would be most likely to approach to help solve these kinds of problems."

Second, we report in Figure 2 the degree distribution across the four types of networks, as well as in the union network. Finally, since these networks are constructed using a name generator, top coding (i.e. naming the maximum number of respondents allowed by the name generator) may be an issue, because it may artificially truncate the degree distribution. In Table 2, we report, for each of our four networks, the percentage of respondents that reported the maximum number of 5 alters. The table shows that the prevalence of top coding is low: it affects about one quarter repondents in the friendship network, and less than one fifth in the family network, and about one tenth in the remaining two networks. Furthermore, for all but

Network	Sample	High uptake	Low uptake	Δ
family	0.18	0.18	0.19	-0.01
friend	0.26	0.28	0.23	0.06
lender	0.13	0.14	0.11	0.03
solver	0.09	0.11	0.07	0.04*

Table 2: **Top coding in network ties elicitation.** This table report the percentage of respondents that reported 5 alters by type of ties. Difference in means are reported in the Δ column, with standard errors clustered at the village level; *p<0.1; **p<0.05; ***p<0.01. Top coding (i.e. reporting 5 alters) has a relatively low prevalence. It affects about one quarter repondents in the friendship network, and less than one fifth in all other networks. Even though high-uptake villages show more top coding than low-uptake villages, the difference is not significant except for the solver network, where it is only significant at the 10 percent level.

one type of network, top coding is not significantly more prevalent in high- than in low-uptake villages. Overall, this suggests that top-coding should not affect results much.

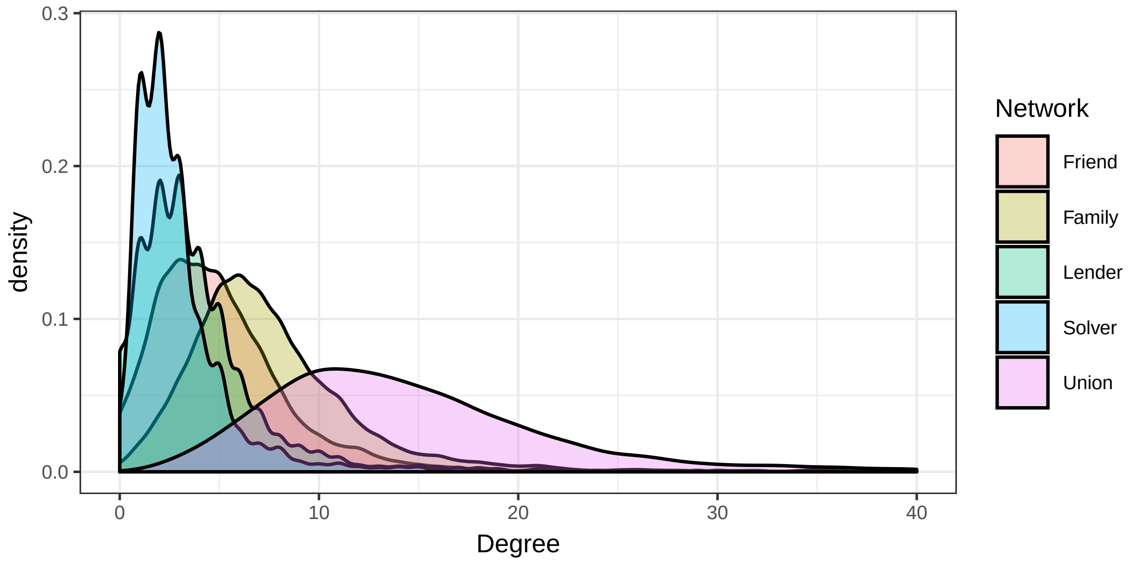

Figure 2: Degree distribution by network type.

2.5 Handling missing network data

Understandably, we were unable to interview every individual in the village. This means there are villagers for whom we only observe a fraction of their network: they were mentioned as ties by other respondents, but were not interviewed in-person. About 18% of named individuals fall in this category. Following standard practice, we exclude those nodes from the analysis. Table 3 reports the number of individuals we surveyed in each village, the number of individuals mentioned by at least one person (“alters”), and the number of adults living in each village, according to 2014 census data. This information allows calculating the number of missing nodes.

Village	N interviewed	N alters	Adult population	Pct. non-interviewed alters	Pct. non-interviewed population	In-degree interviewed	In-degree non-interviewed	Δ
High-uptake								
A	160	216	161	0.26	0.01	84.35	3.86	80.49***
B	30	41	31	0.27	0.03	18.73	2.45	16.28***
C	237	325	295	0.27	0.20	121.77	2.05	119.72***
D	163	212	203	0.23	0.20	85.61	4.35	81.26***
E	263	381	385	0.31	0.32	136.41	4.74	131.68***
F	225	291	198	0.23	-0.14	117.21	3.47	113.74***
G	283	372	320	0.24	0.12	147.25	5.73	141.52***
H	205	296	258	0.31	0.21	107.05	2.74	104.31***
I	254	321	315	0.21	0.19	132.22	3.64	128.57***
Low-uptake								
J	204	281	266	0.27	0.23	106.72	3.92	102.79***
K	192	276	233	0.30	0.18	100.41	4.21	96.19***
L	168	306	230	0.45	0.27	88.49	3.01	85.48***
M	185	264	285	0.30	0.35	98.02	3.8	94.22***
N	229	307	242	0.25	0.05	119.27	3.19	116.08***
O	197	279	274	0.29	0.28	102.19	3.34	98.85***
P	189	262	183	0.28	-0.03	98.99	3.58	95.41***
Pooled								
High-uptake	1820	2455	2166	0.26	0.16	118.86	3.82	115.04***
Low-uptake	1364	1975	1713	0.31	0.20	102.79	3.53	99.26***
All	3184	4430	3879	0.28	0.18	111.97	3.68	108.3***

Table 3: **Network sampling.** N alters reports the number of individuals mentioned as alters in the network survey. Adult population from 2014 census data. The Δ column reports the difference in mean in-degree between interviewed and non-interviewed individuals. Standard errors are heteroskedastic-robust for within-village differences and clustered by village for across-village differences; *p<0.1; **p<0.05; ***p<0.01.

Using the census data, we are able to examine whether we are particularly likely to miss certain types of people, which could have implications for the interpretation of our results. As shown in Table 3 and 4, we find that our sample underrepresents males, and also that the in-degree of those we did not interview is lower than those we interviewed. Together, we show that by systematically missing men, who are more responsive to peer effects (Table 5), we are likely underestimating peer effects.

Uptake	Variable	Sample mean	Census mean	Δ
Low	% female	0.60	0.54	0.05^{***}
Low	Age	37.30	36.38	0.92
Low	secondary education	0.19	0.20	-0.01
High	% female	0.56	0.53	0.04^{***}
High	Age	37.45	37.09	0.36
High	secondary education	0.26	0.29	-0.03
All	% female	0.58	0.54	0.04^{***}
All	Age	37.39	36.78	0.61
All	secondary education	0.23	0.25	-0.02

Table 4: **Comparing sample to census.** Sample mean and census mean for gender, age and secondary education broken down by uptake and for all villages. The sample overrepresents females. Standard errors clustered at the village level; *p<0.1; **p<0.05; ***p<0.01.

	Dependent variable: adopt			
	(1)	(2)	(3)	(4)
# adopting neighbors	0.031^{***}	0.041^{***}	0.015^{***}	0.030^{***}
	(0.011)	(0.010)	(0.005)	(0.004)
degree	0.001^{**}	0.002^{**}	0.001^{**}	0.001^{*}
	(0.001)	(0.001)	(0.001)	(0.001)
age	0.0001	0.0002	−0.0003	−0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
secondary education	−0.002	0.042^{***}	−0.007	0.045^{***}
	(0.017)	(0.010)	(0.017)	(0.010)
female	0.026^{***}	0.016^{**}	0.012	0.033^{***}
	(0.010)	(0.008)	(0.008)	(0.009)
age × # adopting neighbors	−0.0003	-0.0004^{*}		
	(0.0002)	(0.0002)		
secondary education × # adopting neighbors	0.024^{***}		0.026^{***}	
	(0.006)		(0.006)	
female × # adopting neighbors	-0.010^{*}			-0.013^{**}
	(0.005)			(0.005)
Constant	0.163^{***}	0.162^{***}	0.197^{***}	0.169^{***}
	(0.052)	(0.062)	(0.060)	(0.065)
Controls	✓	✓	✓	✓
Observations	3,019	3,019	3,019	3,019
R^2	0.287	0.277	0.284	0.277
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 5: **Sensitivity to peer-effects for sub-populations.** This table reproduces model (2) from Table 2 in the main text but adds interaction terms for age, education, and gender. Males and educated individuals are more sensitive to peer effects. Older individuals are marginally less sensitive to peer effects.

3 Additional information on the puzzle

In this section, we provide additional evidence that supports the puzzle established in section 3 of the paper.

3.1 Heterogenous demand

Variable	Sample	High uptake	Low uptake	Δ	Std. diff.
Public goods					
Public goods summary index	0	0.09	-0.12	0.22	0.39
Working electricity grid in village	0.31	0.56	0	0.56**	1.58*
Road accessible during all seasons	0.73	0.67	0.83	-0.17	0.39
Nursery school in village	0.19	0.33	0	0.33*	1
Government primary school in village	0.5	0.44	0.57	-0.13	0.26
Government health facility in village	0.06	0.11	0	0.11	0.5
Number of functional water sources	1.62	2	1.14	0.86	0.75
Number of functional public toilets/latrines	3.88	5.56	1.71	3.84	0.59
Local goods					
Local goods summary index	0	0.28	-0.37	0.65**	1.63*
Number of savings/community savings groups	1.75	2.22	1.14	1.08	0.84
Number of functional farmers groups and cooperatives	0.94	0.67	1.29	-0.62	0.57
Community center (with physical structure)	0.06	0.11	0	0.11	0.5
Community bar (drinking establishment)	0.19	0.33	0	0.33*	1
General market located within village	0.25	0.33	0.14	0.19	0.46
Market place for crops in village	0.06	0.11	0	0.11	0.5
Community playing field within village	0.38	0.56	0.14	0.41*	0.96
Number of community drying spaces	2.19	3	1.14	1.86*	0.89
Community instruments (e.g. musical, kitchenware)	0.38	0.56	0.14	0.41*	0.96
Number of community bicycles	1.25	1.78	0.57	1.21	0.84
Number of churches in village	0.88	1	0.71	0.29	0.45

Table 6: **Balance table: public and local goods.** High and low-uptake villages largely show no differences in their stock of public and local goods. The only differences point towards high-uptake villages having more goods than low uptake villages. The finding is inconsistent with a possible alternative explanation, according to which high-uptake villages have greater demand for public services because they have less such services. Source: an original survey implemented by the research team capturing the stock of public goods and infrastructure in each of the 16 study villages. Difference in means are reported in the Δ column and tested using a t-test, with heteroskedastic-robust standard errors; *p<0.1; **p<0.05; ***p<0.01. Due to small sample sizes, we also report standardized differences, and denote by * that the 95 percent confidence interval around a standardized difference does not include 0.25.

	High-uptake			Low-uptake			
	$t=0$	$t=1$	(2) - (1)	$t=0$	$t=1$	(5) - (4)	(3) - (6)
Variable	(1)	(2)	(3)	(4)	(5)	(6)	Δ
Teacher absenteeism	0.72	0.7	-0.03	0.72	0.72	0	0.02
Students per class	130.22	95.52	-34.71	121.71	65.02	-56.69**	-21.98
Student teacher ratio	0.02	0.02	0	0.02	0.02	0.01*	0
Student uniform ratio	28.25	30.54	2.29	26.36	21.32	-5.04	-7.32
Student book ratio	60.44	83.79	23.35	52.21	51.79	-0.43	-23.78
Student material ratio	49.88	82.82	32.95	42.21	48.25	6.04	-26.91
PLE pass rate	0.89	_	_	0.9	_	_	_
PLE Grade 1 %	0	_	_	0	_	_	_
PLE Grade 2 %	0.3	_	_	0.42	_	_	_
N schools	8	8	0	8	8	0	0

Table 7: **Balance table: school performance.** High- and low-uptake villages show comparable improvement in the quality of education services. The only difference points towards low-uptake villages seeing more improvement than high-uptake villages. The null difference between high and low-uptake villages is inconsistent with a possible alternative explanation, according to which high-uptake villages have greater demand for public services, such as education (the highest priority sector among message senders). Source: unannounced audits undertaken by the research team; PLE scores are from Arua's district education office. The $t=0$ and $t=1$ columns report mean values for the 8 villages in the high- and low-uptake groups at the baseline and the endline, respectively. Columns 3 and 6 report the difference between the endline and the baseline. Column 7 reports the difference in differences. Difference in means are tested using a t-test, with heteroskedastic-robust standard errors; *p<0.1; **p<0.05; ***p<0.01.

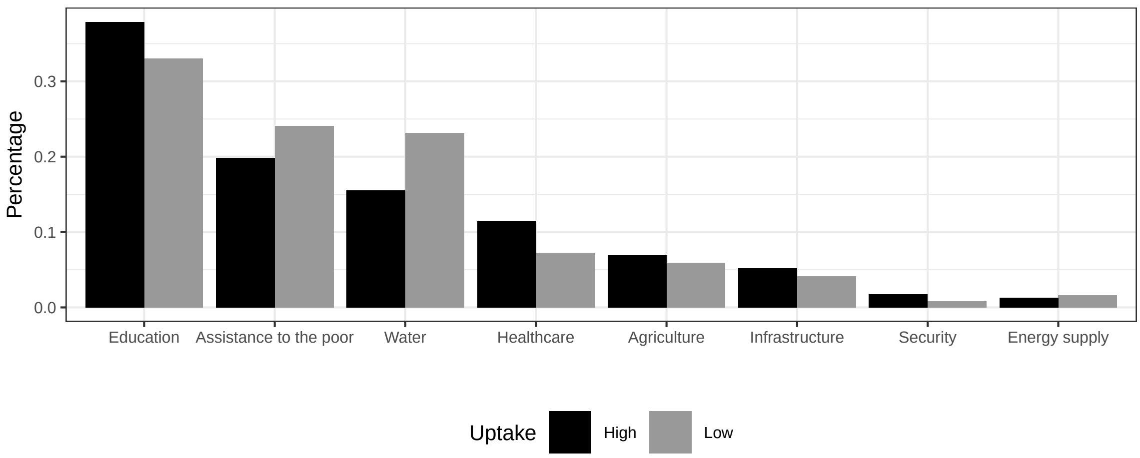

Figure 3: **Policy priorities across high- and low-uptake villages for all surveyed individuals.** High- and low-uptake villages have comparable policy priorities: their orderings of priorities are similar. The finding is inconsistent with an alternative explanation according to which high-uptake villages have greater demand for public services.

3.2 Coordination failure

We conduct a permutation test to check the extent to which individuals agree on the top policy priority that government needs to address. Our survey asks individuals to select one out of nine different priorities. We randomly shuffle those survey responses across the sample and, for each permutation, compute the percentage of pairs of individuals that selected the same priority. We break down this percentage in three categories: individuals that do not live in the same village, individuals that live in the same village but are not connected on the union network, and individuals that live in the same village and are connected on the union network. We then compare the distribution of these three percentages obtained from 10,000 simulations to the values observed in the real data.

Figure 4 shows that individuals agree largely more than predicted by the null distribution, which implies that the divergence between high- and low-uptake villages unlikely owes to coordination failures. On average, two randomly chosen individuals have a probability of agreeing of about 22 percent under the null distribution; irrespective of whether they belong to different villages, are in the same village, or are connected. In the real data, the probability of agreeing is also of about 22 percent for individuals belonging to different villages. It is significantly higher for individuals that belong to the same village (24 percent; a 10 percent increase), and even higher for individuals that are connected (27 percent; a 20 percent increase).

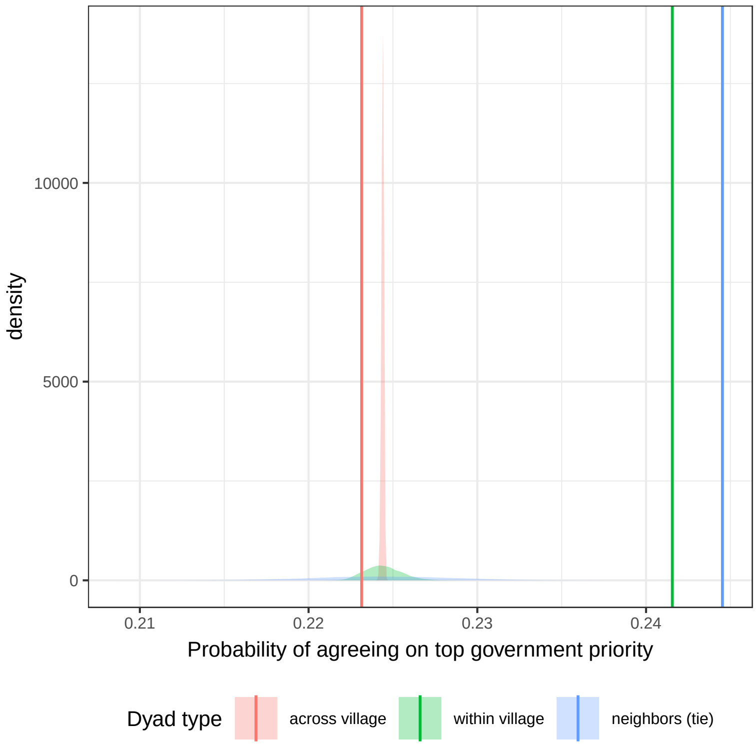

Figure 4: **Agreement on government priorities.** Probability that two individuals agree if they live in different villages, in the same village, or are connected on the union network. Bars represent observed values. Shaded areas represent the distribution of these statistics in a permutation test where government priorities are randomly permuted (10,000 permutations). Individuals that are more closely related agree significantly more than in the null model.

3.3 Private vs. public goods

We coded all relevant incoming messages by *level* of request: private (e.g., personal payments/ assistance unrelated to government program), village (message cites specific service point; our school, the local clinic, etc.), or district. This coding speaks to how valence messages are. We find that the overwhelming majority of messages point out problems at the district or village level, and not at the personal level. The finding is inconsistent with an alternative explanation in which citizens used the platform to request personal favors from their representatives (Figure 5).

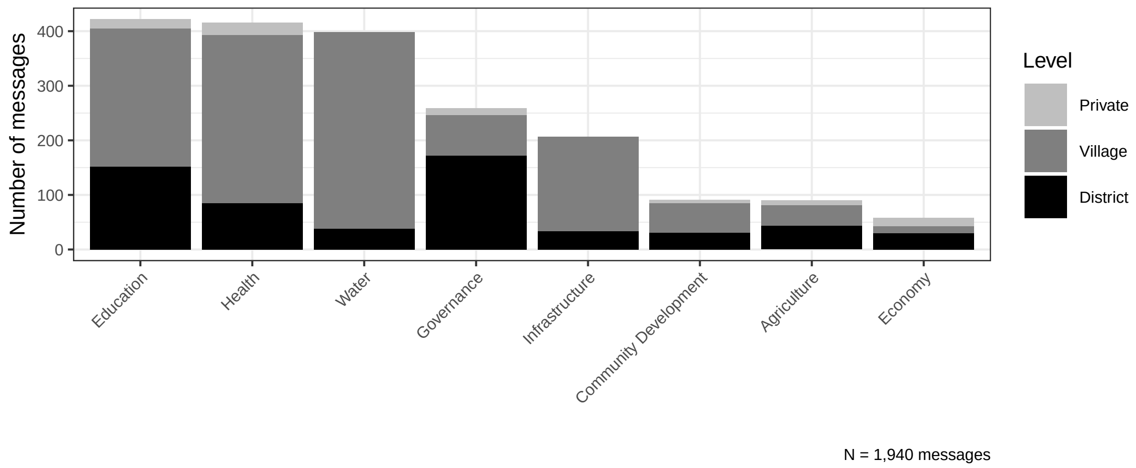

Figure 5: **Messages by topic and level.** Irrespective of the topic, the overwhelming majority of messages point out problems at the district or village level, instead of private problems.

3.4 Government responsiveness

Variable	Sample	High uptake	Low uptake	Δ	Std. diff.
LC5 Chair Turnout of Registered	0.27	0.32	0.21	0.11^{**}	1.17
LC5 Chair Turnout	0.24	0.29	0.18	0.11^{**}	1.55^{*}
Share LC5 winner	0.68	0.61	0.76	-0.15^{**}	1.09

Table 8: **Balance table: voting behavior, 2016 elections.** While high-uptake villages exhibit higher levels of turnout, they are also *less* likely to vote for the incumbent district chairperson. This finding is inconsistent with a possible alternative explanation, according to which high-uptake villages have greater expected benefits because of their level of incumbent support. Source: Uganda's electoral commission. Difference in means are tested using a t-test, with heteroskedastic-robust standard errors; *p<0.1; **p<0.05; ***p<0.01. Due to small sample sizes, we also report standardized differences, and denote by * that the 95 percent confidence interval around a standardized difference does not include 0.25.

3.5 Different seeds

Past work has highlighted the importance—for diffusion of information across networks—of the identity (Banerjee et al., 2013) and network position (Larson, Lewis and Rodrıguez, 2017) of initial 'seeders'. In Table 9, we compare the individual attributes as well as network characteristics of the those attending GAPP's inception meetings and find small and insignificant differences in seeders' characteristics in high- and low-uptake villages.

Variable	Sample	High uptake	Low uptake	Δ	Std. diff.	min	max
Outcome							
% adopters	0.29	0.317	0.224	0.094**		0	1
% satisfied	0.364	0.407	0.222	0.185*		0	1
Individual							
age	40.065	40.452	39.118	1.333		18	88
% females	0.277	0.3	0.224	0.076		0	1
income	2.786	2.785	2.789	-0.005		1	5
secondary education	0.469	0.495	0.408	0.087		0	1
% use phone	0.359	0.376	0.316	0.061		0	1
% leaders	0.282	0.29	0.263	0.027		0	1
political participation index	0.359	0.377	0.316	0.06		-0.878	1.495
pro-sociality	0.2	0.201	0.199	0.002		0	1
Network							
degree	29.427	28.634	31.368	-2.734		3	227
betweenness	0.03	0.027	0.039	-0.013*		0	0.559
clustering coefficient	0.33	0.333	0.323	0.01		0.053	0.844
N							
	262	186	76	110			

Table 9: **Descriptive statistics of meeting attendees in the 16 villages sampled.** Although there are many more meeting attendees in high-uptake villages, attendees in high- and low-uptake villages display comparable characteristics. Difference in means are tested using a t-test with standard errors clustered at the village level; *p<0.1; **p<0.05; ***p<0.01.

Variable	Sample	High uptake	Low uptake	Δ	Std. diff.
Phone use					
% own phone	0.595	0.618	0.565	0.053	
% has used phone for call or text in past month	0.174	0.192	0.15	0.043*	
% has used phone for call in past month	0.616	0.652	0.569	0.083**	
% has used phone for text in past month	0.177	0.196	0.151	0.045*	
Network coverage					
% villages with Airtel coverage	1	1	1	0	n.d.
% villages with MTN coverage	0.94	1	0.86	0.14	0.58
% villages with Orange coverage	1	1	1	0	n.d.

Table 10: **Balance table: phone usage.** Source: phone use data comes from the individual-level survey data analyzed in this paper, network coverage data comes from the survey described in Table 6. High and low-uptake villages differ in the extent to which they use their phones, but do not differ in the extent to which they own phones, or to which service is available. These null differences between high and low-uptake villages are inconsistent with a possible alternative explanation, according to which high-uptake villages have higher opportunity to use the service due to better phone coverage. Difference in means are tested using a t-test, with standard errors clustered at the village level for phone use data, and heteroskedastic-robust standard errors for network coverage data; *p<0.1; **p<0.05; ***p<0.01. Due to small sample sizes, we also report standardized differences for village-level data, and denote by * that the 95 percent confidence interval around a standardized difference does not include 0.25.

4 Model

In this section, we provide further details about the model exposed in section 4 of the paper. The model departs from standard models of technology adoption on networks (e.g. DeGroot, 1974; Bala and Goyal, 1998) in three ways. First, we not only consider goods without externalities, but also goods with positive externalities. Second, we make communication strategic: alters send (possibly inaccurate) reports to ego about their experience with the good. Third, we consider a finite game. This simpler setup reproduces a set of standard results (e.g., that early adopters have higher priors, that goods without externalities benefit from social learning, or that high-degree nodes learn faster[1]). We derive, however, a key additional result: peer effects do not necessarily kick in for goods with externalities.

In what follows, we define additional notation, prove the claims made in the main paper and prove an additional result; namely that high-degree nodes learn faster.

4.1 Additional notation

Similar to S_{i2}, let $S_{i1} \in \mathcal{I}_{i1}(y_{i0})$ be the vector of signals received by i at $t=1$. The vector S_{i1} contains the public signal B_0 and, if $y_{i0}=1$, the private signal s_i. Correspondingly, $\mathcal{I}_{i1}(y_{i0})$ is i's information structure at $t=1$. We have $\mathcal{I}_{i1}(0)=\{0,1\}$, and $\mathcal{I}_{i1}(1)=\mathcal{I}_{i1}(0)\times\{0,1\}$.

4.2 An additional result

Proposition A *Consider equilibrium profile σ_0 with truthful communication, a graph g where there is no tie between agents i and j, and graph g' constructed by adding to g a tie between i and j. We have*

$$V_{ig}^{\sigma_0}(y_{i0},y_{-i0}) \leq V_{ig'}^{\sigma_0}(y_{i0},y_{-i0})$$

Proof of proposition A. Using proposition 1, we know that on g', agent i may receive one more signal at $t=2$ than on g, depending on whether $\pi_j \geq a_{j0}^{\sigma_0}$. Suppose not. Then $V_{ig}^{\sigma}(y_{i0},y_{-i0}) = V_{ig'}^{\sigma_0}(y_{i0},y_{-i0})$. Suppose agent i receives one more signal. Let $\mathcal{I}_{i2}^{g}(y_{i0},y_{-i0})$ be i's information structure at $t=2$ on graph g following actions y_{i0},y_{-i0}, with generic element $S_{i2}^{g} \in \mathcal{I}_{i2}^{g}(y_{i0},y_{-i0})$. Note that in the case under consideration, we have $\mathcal{I}_{i2}^{g'}(y_{i0},y_{-i0}) = \mathcal{I}_{i2}^{g}(y_{i0},y_{-i0})\times\{0,1\}$. Define the function $\varphi : \mathcal{I}_{i2}^{g'}(y_{i0},y_{-i0}) \rightarrow \mathcal{I}_{i2}^{g}(y_{i0},y_{-i0})$ that associates to each $S_{i2}^{g'} \in \mathcal{I}_{i2}^{g'}(y_{i0},y_{-i0})$ the same $S_{i2}^{g'}$ without that additional signal from j, and note that this object belongs to $\mathcal{I}_{i2}^{g}(y_{i0},y_{-i0})$. Define $y_{i2}^{0}(S_{i2}^{g'}) = y_{i2}(\varphi(S_{i2}^{g'}))$. Recall that on g', y_{i2}^{*} solves $\max_{y_{i2}} \mathbb{E}_{\theta}[u_i(y_i,\theta,.)|S_{i2}^{g'}]$. As such, it must be that $\mathbb{E}_{\theta}[u_i(y_{i2}^{*}(S_{i2}^{g'}),\theta,.)|S_{i2}^{g'}] \geq \mathbb{E}_{\theta}[u_i(y_{i2}^{0}(S_{i2}^{g'}),\theta,.)|S_{i2}^{g'}]$. This implies

$$\begin{aligned}
V_{ig'}^{\sigma_0}(y_{i0},y_{-i0}) = & \sum_{S_{i2}^{g'} \in \mathcal{I}_{i2}(y_{i0},y_{-i0})} \mathbb{E}_{\theta}[u(y_{i2}^{*}(S_{i2}^{g'}),\theta,.)|S_{i2}^{g'}] \Pr_{\sigma_0}(S_{i2}^{g'}) \geq \\
& \sum_{S_{i2}^{g'} \in \mathcal{I}_{i2}(y_{i0},y_{-i0})} \mathbb{E}_{\theta}[u(y_{i2}^{0}(S_{i2}^{g'}),\theta,.)|S_{i2}^{g'}] \Pr_{\sigma_0}(S_{i2}^{g'}) = V_{ig}^{\sigma_0}(y_{i0},y_{-i0})
\end{aligned}$$

■

4.3 Proof of propositions in the paper

Proof of proposition 1. We consider the case without externalities first, and show existence of a threshold strategy at $t=2$. Let $\hat{\pi}_{i\sigma} \equiv \Pr_{\sigma}(\theta = H|S_i)$ be i's posterior at $t=2$ under the

[1] See proposition A below.

consistent posterior beliefs implied by equilibrium profile σ. In equilibrium, the action y_{i2}^* maximizes i's expected payoff. This implies

$$y_{i2}^* = 1 \iff \mathbb{E}_\sigma[u(1,\theta)|S_{i2}] \geq \mathbb{E}_\sigma[u(0,\theta)|S_{i2}]$$

Let $\hat{\pi}_{i\sigma}(S_{i2}) \equiv \Pr_\sigma(\theta = H|S_{i2})$ be i's posterior at $t=1$ under profile σ. Then, we have $\mathbb{E}_\sigma[u(1,\theta)|S_{i2}] = \hat{\pi}_{i\sigma}(S_{i2})(p_H - c) + (1-\hat{\pi}_{i\sigma}(S_{i2}))(p_L - c)$, and $\mathbb{E}_\sigma[u(1,\theta)|S_{i2}] = 0$. Solving for $\hat{\pi}_{i\sigma}(S_{i2})$ gives the following inequality:

$$\mathbb{E}_\sigma[u(1,\theta)|S_{i2}] \geq \mathbb{E}_\sigma[u(0,\theta)|S_{i2}] \iff \hat{\pi}_{i\sigma}(S_{i2}) \geq \frac{c-p_L}{p_H-p_L}$$

Note that $\hat{\pi}_{i\sigma}(S_{i2}) = \frac{\pi_i \Pr_\sigma(S_{i2}|H)}{\pi_i \Pr_\sigma(S_{i2}|H) + (1-\pi_i)\Pr(S_{i2}|L)}$. Substituting this expression into the above inequality and rearranging, we get

$$\hat{\pi}_{i\sigma}(S_{i2}) \geq \frac{c-p_L}{p_H-p_L} \iff L_\sigma(S_{i2}) \geq \frac{c-p_L}{p_H-c}\frac{1-\pi_i}{\pi_i} > 0$$

Taking logs, we get

$$y_{i2}^* = 1 \iff l_\sigma(S_{i2}) \geq a_{i2}^\sigma \equiv \log\left[\frac{c-p_L}{p_H-c}\frac{1-\pi_i}{\pi_i}\right]$$

We now show existence of a threshold strategy at $t=0$. In equilibrium, the action y_{i0}^* maximizes i's expected payoff. With M_i^{s*} being i's equilibrium message-sending strategy at $t=1$, this implies

$$y_{i0}^* = 1 \iff \mathbb{E}_\sigma[u(1,\theta) + v(M_i^{s*})^\gamma + u(y_{i2}^*,\theta)^{\gamma^2}|y_{i0}=1] \geq \mathbb{E}_\sigma[u(0,\theta) + v(M_i^{s*})^\gamma + u(y_{i2}^*,\theta)^{\gamma^2}|y_{i0}=0]$$

Note that $\mathbb{E}_\sigma[u(1,\theta)] = \pi_i(p_H - c) + (1-\pi_i)(p_L - c)$, and $\mathbb{E}_\sigma[u(0,\theta)] = 0$. Furthermore, note that $V_{i1}^\sigma(\mathcal{I}_{i1}(y_{i0})) \equiv \mathbb{E}_\sigma[v(M_i^{s*})^\gamma + u(y_{i2}^*,\theta)^{\gamma^2}|y_{i0}]$ is the Blackwell information value of the subgame that starts at $t=1$ following action y_{i0} under profile σ. Substituting and solving for π_i in the above inequality gives

$$y_{i0}^* = 1 \iff \pi_i \geq a_{i0}^\sigma \equiv \frac{[c-p_L] - [V_{i1}^\sigma(\mathcal{I}_{i1}(1)) - V_{i1}^\sigma(\mathcal{I}_{i1}(0))]}{p_H - p_L}$$

We now consider the case with externalities, and proceed similarly. We first show existence of a threshold strategy at $t=2$. In equilibrium, the action y_{i2}^* maximizes i's expected payoff. This implies

$$y_{i2}^* = 1 \iff \mathbb{E}_\sigma[u(1,\theta,n_{-i2})|S_{i2}] \geq \mathbb{E}_\sigma[u(0,\theta,n_{-i2})|S_{i2}]$$

Note that n_{-i2} is a random variable with discrete support $\{0,...,N-1\}$. Since strategy profile σ is consistent, n_{-i2} has a well-defined posterior distribution $\Pr_\sigma(n_{-i2}|S_{i2})$. We have

$$\mathbb{E}_\sigma[u(1,\theta,n_{-i2})|S_{i2}] = \sum_{n=0}^{N-1}\Pr_\sigma(n_{-i2}=n|S_{i2})[\hat{\pi}_{i\sigma}(S_{i2})q(n+1,H) + (1-\hat{\pi}_{i\sigma}(S_{i2}))q(n+1,L)] - c$$
$$\mathbb{E}_\sigma[u(0,\theta,n_{-i2})|S_{i2}] = \sum_{n=0}^{N-1}\Pr_\sigma(n_{-i2}=n|S_{i2})[\hat{\pi}_{i\sigma}(S_{i2})q(n,H) + (1-\hat{\pi}_{i\sigma}(S_{i2}))q(n,L)]$$

With $\Delta q_\theta(n) \equiv q(n+1,\theta) - q(n,\theta) \geq 0$, we get

$$y_{i2}^* = 1 \iff \sum_{n=0}^{N-1}\Pr_\sigma(n_{-i2}=n|S_{i2})[\hat{\pi}_{i\sigma}(S_{i2})\Delta q_H(n) + (1-\hat{\pi}_{i\sigma}(S_{i2}))\Delta q_L(n)] \geq c$$

Define $C_\sigma(S_{i2}) \equiv \sum_{n=0}^{N-1}\Pr_\sigma(n_{-i2}=n|S_{i2})\Delta q_H(n)$ and $D_\sigma(S_{i2}) \equiv \sum_{n=0}^{N-1}\Pr_\sigma(n_{-i2}=n|S_{i2})\Delta q_L(n)$.

Substituting and rearranging, we get

$$y_{i2}^* = 1 \iff \hat{\pi}_{i\sigma}(S_{i2})[C_\sigma(S_{i2}) - D_\sigma(S_{i2})] \geq c - D_\sigma(S_{i2})$$

Note that since $q(n,L) < c$ for any $n \in \{0,...,N\}$, it must be that $c - D_\sigma(S_{i2}) > 0$. Also note that since $\Delta q_H(n) > \Delta q_L(n)$, it must be that $C_\sigma(S_{i2}) - D_\sigma(S_{i2}) > 0$. If $\frac{c-D_\sigma(S_{i2})}{C_\sigma(S_{i2})-D_\sigma(S_{i2})} \geq 1 \iff c \geq C_\sigma(S_{i2})$, it must be that $y_{i2}^* = 0$ for any $S_{i2} \in \mathcal{I}_{i2}(y_{i0}, y_{-i0})$. Otherwise, $C_\sigma(S_{i2}) > c$, and we can rearrange to get

$$y_{i2}^* = 1 \iff L_\sigma(S_{i2}) \geq \frac{c - D_\sigma(S_{i2})}{C_\sigma(S_{i2}) - c} \frac{1-\pi_i}{\pi_i} > 0$$

Taking logs, we get

$$y_{i2}^* = 1 \iff l_\sigma(S_{i2}) \geq a_{i2}^\sigma(S_{i2}) \equiv \log\left[\frac{c - D_\sigma(S_{i2})}{C_\sigma(S_{i2}) - c} \frac{1-\pi_i}{\pi_i}\right]$$

If $C_\sigma(S_{i2}) \leq c$, then define $a_{i2}^\sigma(S_{i2}) = \max\{l(S_{i2})\} + 1$.

We now show existence of a threshold strategy at $t = 0$. In equilibrium, the action y_{i0}^* maximizes i's expected payoff. This implies

$$\begin{aligned} y_{i0}^* = 1 \iff & \mathbb{E}_\sigma[u(1,\theta,n_{-i0}) + v(M_i^{s*})^\gamma + u(y_{i2}^*,\theta,n_{-i2})^{\gamma^2} | y_{i0} = 1] \geq \\ & \mathbb{E}_\sigma[u(0,\theta,n_{-i0}) + v(M_i^{s*})^\gamma + u(y_{i2}^*,\theta,n_{-i2})^{\gamma^2} | y_{i0} = 0] \end{aligned}$$

Since strategy profile θ has well-defined beliefs $\Pr_\sigma(n_{-i0})$ over n_{-i0}. As such, expected payoffs at $t = 0$ write

$$\mathbb{E}_\sigma[u(1,\theta,n_{-i0})] = \sum_{n=1}^{N-1} \Pr_\sigma(n_{-i0} = n)[\pi_i q(n+1,H) + (1-\pi_i)q(n+1,L)] - c$$

$$\mathbb{E}_\sigma[u(0,\theta,n_{-i0})] = \sum_{n=1}^{N-1} \Pr_\sigma(n_{-i0} = n)[\pi_i q(n,H) + (1-\pi_i)q(n,L)]$$

Furthermore, note that $V_{i1}^\sigma(\mathcal{I}_{i1}(y_{i0})) \equiv \mathbb{E}_\sigma[v(M_i^{s*})^\gamma + u(y_{i2}^*,\theta)^{\gamma^2} | y_{i0}]$ is the Blackwell information value of the subgame that starts at $t = 1$ following action y_{i0}. Also, define $C_\sigma \equiv \sum_{n=1}^{N-1} \Pr_\sigma(n_{-i0} = n)\Delta q_H(n)$ and $D_\sigma \equiv \sum_{n=1}^{N-1} \Pr_\sigma(n_{-i0} = n)\Delta q_L(n)$. Note that $q_L(n) < c$ implies $c > D_\sigma$, and that $\Delta q_H(n) > \Delta q_L(n)$ implies $C_\sigma - D_\sigma > 0$.

Substituting and solving for π_i in the above inequality gives

$$y_{i0}^* = 1 \iff \pi_i \geq a_{i0}^\sigma \equiv \frac{[c - D_\sigma] - [V_{i1}^\sigma(\mathcal{I}_{i1}(1)) - V_{i1}^\sigma(\mathcal{I}_{i1}(0))]}{C_\sigma - D_\sigma}$$

■

Proof of proposition 2. This is an immediate application of Blackwell et al.'s (1951) theorem. Under σ, with some probability, we have $m_{ji} \neq s_j$. Conversely, under σ_0, we always have $m_{ji} = s_j$. As such, one may *garble* the conditional distribution of m_{ji} under σ_0 to recover the conditional distribution of m_{ji} under σ. Blackwell's information theorem therefore implies $V_{ig}^\sigma(y_{i0}, y_{-i0}) \leq V_{ig}^{\sigma_0}(y_{i0}, y_{-i0})$. ■

Proof of proposition 3. Consider the action of agent i at $t = 1$ under some consistent strategy profile σ and suppose that $y_{i0} = 1$. Let $\mathcal{M}_i \equiv \{0,1\}^{|N_{ig}|}$ be the set of messages i can send to her neighbors, and $\mathcal{L}_{\mathcal{M}_i}$ be the set of lotteries over $\mathcal{M}_i$. Recall that in equilibrium, the lottery $l \in \mathcal{L}_{\mathcal{M}_i}$ solves $\max_{l \in \mathcal{L}_{\mathcal{M}_i}} \mathbb{E}_l[v(m)] + \mathbb{E}_\sigma[u(y_{i2},\theta)^\gamma | S_{i1}]$. Consider the lotteries l_0 where i is truthful, and l_k where i lies with some probability. We have

$$[\mathbb{E}_{l_k}[v(m)] + \mathbb{E}_\sigma[u(y_{i2},\theta)^\gamma | S_{i1}]] - [\mathbb{E}_{l_0}[v(m)] + \mathbb{E}_\sigma[u(y_{i2},\theta)^\gamma | S_{i1}]] = -k\kappa,$$

for any y_{i2}. When $\kappa = 0$, i is indifferent between telling the truth and lying. Truthful communication is therefore an equilibrium. When $\kappa > 0$, i strictly prefers to tell the truth to lying, so truthful communication is the unique perfect Bayesian equilibrium. ■

Proof of proposition 4. Consider the action of agent i at $t = 1$ under some consistent strategy profile σ and suppose that $y_{i0} = 1$. Let $\mathcal{M}_i \equiv \{0,1\}^{|N_{ig}|}$ be the set of messages i can send to her neighbors, and $\mathcal{L}_{\mathcal{M}_i}$ be the set of lotteries over $\mathcal{M}_i$. Recall that in equilibrium, the lottery $l \in \mathcal{L}_{\mathcal{M}_i}$ solves $\max_{l \in \mathcal{L}_{\mathcal{M}_i}} \mathbb{E}_l[v(m)] + \mathbb{E}_\sigma[u(y_{i2}, \theta, y_{-i2})^\gamma | S_{i1}, l]$.

Let σ_0 be a consistent strategy profile with truthful communication. Let $m_0 = \{s_i\}^{|N_{ig}|} \in \mathcal{M}_i$ be the truthful message. In profile σ_0, agent i plays lottery $l_0 \in \mathcal{L}_{\mathcal{M}_i}$ such that $\Pr(m = m_0) = 1$. The profile σ_0 is an equilibrium if and only if

$$[\mathbb{E}_l[v(m)] + \mathbb{E}_{\sigma_0}[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, l]] - [v(m_0) + \mathbb{E}_{\sigma_0}[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, m_0]] \leq 0 \qquad (1)$$

for any $i \in N, S_{i1} \in \mathcal{I}_{i1}(1), l \in \mathcal{L}_{\mathcal{M}_i} \setminus \{l_0\}$.

Let $p_{lm} \in [0,1]$ be the probability of sending the messages $m \in \mathcal{M}_i$ under lottery l, and $k_m \geq 0$ the number of lies associated with messages m. Note that one can break down the payoff associated with lottery l as follows:

$$\mathbb{E}_l[v(m)] + \mathbb{E}_\sigma[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, l] =$$
$$\sum_{m \in \mathcal{M}_i \setminus \{m_0\}} p_{lm}(\mathbb{E}_\sigma[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, m] - k_m \kappa) + p_{l0} \mathbb{E}_\sigma[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, m_0]$$

Using this decomposition as well as the fact that $\sum_{m \in \mathcal{M}_i \setminus \{m_0\}} p_{lm} = 1 - p_{lm_0}$ on equation 1 we get after rearranging that profile σ_0 is an equilibrium if and only if

$$\sum_{m \in \mathcal{M}_i \setminus \{m_0\}} p_{lm} \left(\mathbb{E}_{\sigma_0}[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, l] - \mathbb{E}_{\sigma_0}[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, m_0]\right) \leq \kappa \sum_{m \in \mathcal{M}_i \setminus \{m_0\}} p_{lm} k_m$$

for any $i \in N, S_{i1} \in \mathcal{I}_{i1}(1), l \in \mathcal{L}_{\mathcal{M}_i} \setminus \{l_0\}$. Since, for any $l \neq l_0$, we have $\sum_{m \in \mathcal{M}_i \setminus \{m_0\}} p_{lm} > 0$ and $k_m > 0$ for any $m \neq m_0$, truthful communication is an equilibrium if and only if

$$\kappa \geq \tilde{\kappa}_1 \equiv \max_{i \in N, S_{i1} \in \mathcal{I}_{i1}(1), l \in \mathcal{L}_{\mathcal{M}_i} \setminus \{l_0\}} \frac{\sum_{m \in \mathcal{M}_i \setminus \{m_0\}} p_{lm} \left(\mathbb{E}_{\sigma_0}[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, l] - \mathbb{E}_{\sigma_0}[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, m_0]\right)}{\sum_{m \in \mathcal{M}_i \setminus \{m_0\}} p_{lm} k_m}$$

Conversely, consider the set Σ of consistent strategy profiles with some lying. Let $l_{i\sigma} \in \mathcal{M}_i$ be the lottery that i plays under profile $\sigma \in \Sigma$. If we have, for any $\sigma \in \Sigma, i \in N, S_{i1} \in \mathcal{I}_{i1}(1)$, that

$$[\mathbb{E}_{l_\sigma}[v(m)] + \mathbb{E}_\sigma[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, l_\sigma]] - [v(m_0) + \mathbb{E}_{\sigma_0}[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, m_0]] < 0,$$

then lying is not an equilibrium. Using the same approach as with equation 1, this yields that lying is not an equilibrium if

$$\kappa > \tilde{\kappa}_2 \equiv \max_{\sigma \in \Sigma, i \in N, S_{i1} \in \mathcal{I}_{i1}(1)} \frac{\sum_{m \in \mathcal{M}_i \setminus \{m_0\}} p_{lm} \left(\mathbb{E}_\sigma[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, l] - \mathbb{E}_\sigma[u(y_{i2}, \theta, n_{-i2})^\gamma | S_{i1}, m_0]\right)}{\sum_{m \in \mathcal{M}_i \setminus \{m_0\}} p_{lm} k_m}$$

Since $u(y_{i2}, \theta, n_{-i2}) \leq 1$ and $k_m \geq 1$ for any $m \in \mathcal{M}_i \setminus \{m_0\}$, it must be that if $\tilde{\kappa}_2$ exists, then $\tilde{\kappa}_2 \leq 1$. If $\tilde{\kappa}_2$ exists, define $\kappa_2 \equiv \tilde{\kappa}_2$. Otherwise, define $\kappa_2 \equiv 1$. We show similarly that if $\tilde{\kappa_1}$ exists, then $\tilde{\kappa}_1 \leq 1$. If $\tilde{\kappa_1}$ exists, then define $\kappa_1 \equiv \max\{0, \tilde{\kappa}_1\}$. Otherwise, define $\kappa_1 = 1$.

Finally, it must be that $\kappa_1 \leq \kappa_2$ for otherwise, the game has no equilibrium for any $\kappa \in (\kappa_2, \kappa_1)$. Yet, we are considering a finite game of incomplete information. As such, this game always has a perfect Bayesian equilibrium.

■

Lemma A *Let $\mathcal{S} \equiv (s_1, ..., s_K)$ be a vector of K independent binary random variables such that* $\Pr(s_k = 1) \equiv p_k$ *for any $k \in \{1, ..., K\}$, and $A \equiv (a_1, ..., a_K)$ a vector of K coefficients such that $a_k \neq 0$ for any $k \in \{1, ..., K\}$. For any non-empty set $S \subseteq \mathcal{S}$, define $f(S) = a_S + \sum_{k=1}^{K} 1\{s_k \in S\} a_k s_k$. Consider the non-empty subsets of $\mathcal{S}$, S_1 and S_2 such that $S_2 \setminus S_1 \neq \emptyset$, and $S_1' = S_1 \cup \{s_l\}$, with $s_l \in S_2 \setminus S_1$. We have*

$$\rho[f(S_1), f(S_2)] \leq \rho[f(S_1'), f(S_2)]$$

Proof of lemma A. Recall that $\rho(x,y) = \frac{\text{Cov}(x,y)}{\sqrt{\mathbb{V}(x)\mathbb{V}(y)}}$, with $\text{Cov}(x,y) = \mathbb{E}(xy) - \mathbb{E}(x)\mathbb{E}(y)$ the covariance of x and y, and note that

$$\begin{aligned}
f(S)^2 =& a_S^2 + \sum_{k=1}^{K} 1\{s_k \in S\} a_k^2 s_k^2 + \sum_{k=1}^{K} \sum_{k' \neq k} 1\{s_k \in S\} 1\{s_{k'} \in S\} a_k a_{k'} s_k s_{k'} + \\
& 2a_S \sum_{k=1}^{K} 1\{s_k \in S\} a_k s_k \\
f(S_1)f(S_2) =& a_{S_1} a_{S_2} + \sum_{k=1}^{K} 1\{s_k \in S_1\} 1\{s_k \in S_2\} a_k^2 s_k^2 + \sum_{k=1}^{K} \sum_{k' \neq k} 1\{s_k \in S_1\} 1\{s_{k'} \in S_2\} a_k a_{k'} s_k s_{k'} + \\
& a_{S_2} \sum_{k=1}^{K} 1\{s_k \in S_1\} a_k s_k + a_{S_1} \sum_{k=1}^{K} 1\{s_k \in S_2\} a_k s_k
\end{aligned}$$

From this, we get

$$\begin{aligned}
\mathbb{V}[f(S)] &= \sum_{k=1}^{K} 1\{s_k \in S\} a_k^2 p_k (1 - p_k) \\
\text{Cov}[f(S_1), f(S_2)] &= \sum_{k=1}^{K} 1\{s_k \in S_1\} 1\{s_k \in S_2\} a_k^2 p_k (1 - p_k)
\end{aligned}$$

Let $\lambda \equiv a_l^2 p_l (1 - p_l)$. Using the above, we get $\mathbb{V}[f(S_1')] = \mathbb{V}[f(S_1)] + \lambda$ and $\text{Cov}[f(S_1', S_2)] = \text{Cov}[f(S_1), f(S_2)] + \lambda$. Therefore,

$$\begin{aligned}
\rho[f(S_1'), f(S_2)] - \rho[f(S_1), f(S_2)] &= \frac{\text{Cov}[f(S_1', S_2)]}{\sqrt{\mathbb{V}[f(S_1')]\mathbb{V}[f(S_2)]}} - \frac{\text{Cov}[f(S_1, S_2)]}{\sqrt{\mathbb{V}[f(S_1)]\mathbb{V}[f(S_2)]}} \\
&\propto \mathbb{V}[f(S_1)]\text{Cov}[f(S_1', S_2)] - \mathbb{V}[f(S_1')]\text{Cov}[f(S_1, S_2)] \\
&= \lambda \left(\mathbb{V}[f(S_1)] - \text{Cov}[f(S_1, S_2)] \right) \\
&= \lambda \sum_{k=1}^{K} 1\{s_k \in S_1\} 1\{s_k \notin S_2\} a_k^2 p_k (1 - p_k) \geq 0
\end{aligned}$$

■

Proof of proposition 5. Using proposition 1, we know that on g, agent i receives $k_{ig} \equiv 1 + |\{j \in N_i(g) : \pi_j \geq a_{j0}^{\sigma_0}\}| + 1\{\pi_i \geq a_{i0}^{\sigma_0}\} \geq 0$ signals at $t = 2$; that is, the public signal B_0, the messages sent by her adopting neighbors, and her private signal (if she adopts). Similarly, j receives $k_{jg} \geq 0$ signals at $t = 2$. On g', agent i may receive one more signal, depending on whether j adopts or not at $t = 0$: $k_{ig'} = k_{ig} + 1\{\pi_j \geq a_{j0}^{\sigma_0}\}$. Similarly, we have $k_{jg'} = k_{jg} + 1\{\pi_i \geq a_{i0}^{\sigma_0}\}$. This gives us three cases:

1. $k_{ig'} = k_{ig}$ and $k_{jg'} = k_{jg}$
2. $k_{ig'} = k_{ig} + 1$ and $k_{jg'} = k_{jg}$ (without loss of generality)

3. $k_{ig'} = k_{ig} + 1$ and $k_{jg'} = k_{jg} + 1$

In case 1, we have $S_{i2}^{g'} = S_{i2}^{g}$ and $S_{j2}^{g'} = S_{j2}^{g}$. As such, we have $\rho[l_{\sigma_0}(S_{i2}^{g}), l_{\sigma_0}(S_{j2}^{g})] = \rho[l_{\sigma_0}(S_{i2}^{g'}), l_{\sigma_0}(S_{j2}^{g'})]$.

In case 2, we have $S_{i2}^{g'} = S_{i2}^{g} \cup \{s_j\}$ and $S_{j2}^{g'} = S_{j2}^{g}$. Note that

$$L_{\sigma_0}(S_{i2}^{g}) = \frac{q(H,.)^{B_0}[1-q(H,.)]^{1-B_0}\prod_{j\in N} r_H^{1\{s_j \in S_{i2}^{g}\}s_j}(1-r_H)^{1\{s_j \in S_{i2}^{g}\}(1-s_j)}}{q(L,.)^{B_0}[1-q(L,.)]^{1-B_0}\prod_{j\in N} r_L^{1\{s_j \in S_{i2}^{g}\}s_j}(1-r_L)^{1\{s_j \in S_{i2}^{g}\}(1-s_j)}}$$

Taking logs, we get

$$l_{\sigma_0}(S_{i2}^{g}) = \left(\log\left[\frac{1-q(H,.)}{1-q(L,.)}\right] + \sum_{j\in N} 1\{s_j \in S_{i2}^{g}\}\log\left[\frac{1-r_H}{1-r_L}\right]\right) + B_0\log\left[\frac{q(H,.)[1-q(L,.)]}{q(L,.)[1-q(H,.)]}\right] + \sum_{j\in N} 1\{s_j \in S_{i2}^{g}\}s_j\log\left[\frac{r_H(1-r_L)}{r_L(1-r_H)}\right]$$

Using lemma A, we get $\rho[l_{\sigma_0}(S_{i2}^{g}), l_{\sigma_0}(S_{j2}^{g})] \leq \rho[l_{\sigma_0}(S_{i2}^{g'}), l_{\sigma_0}(S_{j2}^{g'})]$.

In case 3, we have $S_{i2}^{g'} = S_{i2}^{g} \cup \{s_j\}$ and $S_{j2}^{g'} = S_{j2}^{g} \cup \{s_i\}$. We first use the same argument as in case 2 and get that $\rho[l_{\sigma_0}(S_{i2}^{g}), l_{\sigma_0}(S_{j2}^{g})] \leq \rho[l_{\sigma_0}(S_{i2}^{g'}), l_{\sigma_0}(S_{j2}^{g})]$. Using lemma A again, we get that $\rho[l_{\sigma_0}(S_{i2}^{g'}), l_{\sigma_0}(S_{j2}^{g})] \leq \rho[l_{\sigma_0}(S_{i2}^{g'}), l_{\sigma_0}(S_{j2}^{g'})]$. ■

5 Robustness checks and causal inference

In this section, we describe in greater detail the robustness checks and further analyses reported in Section 5 of the paper. We first discuss our robustness checks, then the analyzes we conduct to show that are estimates likely lend themselves to a causal interpretation.

5.1 Robustness checks

5.1.1 Random effects model

Our first robustness check estimates a multilevel specification to ascertain that the difference in peer effects between high- and low-uptake villages does not owe to outlying villages. We estimate a Bayesian multilevel model using the `rstanarm` package with the `R` statistical software. Our priors use the `rstanarm` defaults; that is, weakly informative normal priors for the coefficients, with mean 0 and standard deviation 10 for the intercept and 2.5 for other coefficients, and an exponential prior for the standard deviation of the error term. Each scale is then multiplied by the by the standard deviation of the dependent variable. Similarly, we use the correlation matrix decomposition of the covariance matrix, with and LKJ prior for the correlation matrix (regularization parameter of 1), a Dirichlet prior for the simplex vector (concentration parameter of 1), and a Gamma prior for the scale parameter (shape and scale of 1). We estimate those models using `rstan`'s NUTS sampler with 4 MCMC chains of 10,000 draws each, with a burn-in of 5,000 draws. Table 11 reports the main models, which use random intercepts by village instead of village-level fixed effects. While model (1) reproduces our main specification (Model 2, Table 2 in the paper) using random intercepts, the model 2 adds random slopes on the number of adopting neighbors by village. We find that our conclusion of differential peer effects does apply to virtually all villages: peer effects are positive and significantly different from zero in all high-uptake villages (A to I) but one (village F). They are not significantly different from 0 in low-uptake villages (J to P).

	Dependent variable: adopt	
	(1)	(2)
# adopting neighbors (β_1)	0.002	
	(−0.009, 0.014)	
# adopting neighbors × high-uptake (β_2)	0.026	
	(0.014, 0.038)	
# adopting neighbors (village A)		0.040
		(0.025, 0.055)
# adopting neighbors (village C)		0.024
		(0.012, 0.037)
# adopting neighbors (village D)		0.022
		(0.006, 0.038)
# adopting neighbors (village E)		0.022
		(0.012, 0.033)
# adopting neighbors (village F)		−0.010
		(−0.035, 0.016)
# adopting neighbors (village G)		0.030
		(0.020, 0.040)
# adopting neighbors (village H)		0.024
		(0.012, 0.037)
# adopting neighbors (village I)		0.023
		(0.008, 0.039)
# adopting neighbors (village J)		0.005
		(−0.019, 0.031)
# adopting neighbors (village K)		−0.013
		(−0.053, 0.025)
# adopting neighbors (village L)		0.009
		(−0.010, 0.029)
# adopting neighbors (village M)		−0.001
		(−0.016, 0.015)
# adopting neighbors (village N)		−0.006
		(−0.032, 0.019)
# adopting neighbors (village O)		−0.001
		(−0.032, 0.032)
# adopting neighbors (village P)		−0.008
		(−0.035, 0.020)
degree	0.001	0.002
	(0.001, 0.002)	(0.001, 0.002)
Controls	✓	✓
Random intercept	✓	✓
Random slope	—	✓
$\beta_1 + \beta_2$	0.028	
	(0.022, 0.035)	
Observations	2,991	2,991
Note:	95 percent credible intervals in parenthesis.	

Table 11: **Main random effect models.** Model 1 reproduces our main specification (Model 2, Table 2 in the paper) with random effects. Model 2 use random slopes for the number of adopting and hearing neighbors respectively. Peer effects are positive and significantly different from zero in all high-uptake villages (A to I) but one (village F). They are not significantly different from 0 in low-uptake villages (J to P).

5.1.2 Alternative specifications

We then explore several alternative specifications. First, throughout the paper, we defined as an adopter any individual that has used the platform at least once in the past 12 months. Here, we consider stronger definitions, and define as adopters individuals that have used the platform at least 3, or 5 times in the past 12 months. Results are robust to using stronger definitions (Table 12).

Second, we explore whether results are sensitive to using Logistic regression instead of a linear probability model. Results reported in Table 13 suggest that peer effects are significant, on average, irrespective of this choice.

Third, we test whether our main results are sensitive to dropping village B, which has a significantly smaller number of respondents (30) compared to the other villages (mean number of respondents is 210). Table 14 shows that results are virtually unchanged. We find a strong positive relationship between the number (or share) of adopting neighbors and one's adoption choice, with magnitudes almost identical to the main specification.

Fourth, we test whether our main results are sensitive to using undirected ties, which may capture a different notion of influence. Table 15 amends our main specification (Table 2 in the main text) but uses the directed union network instead of the undirected union network. The undirected union network is the union of the undirected friend, family, lender, and solver networks. In other words, there is a tie from i to j in this network if there is a tie from i to j in any of those four networks. Since the number of out-neighbors is capped at 5 by design, we focus on in-neighbors: the nodes j that have a tie to i. We count the number of adopting in-neighbors. Similarly, we consider in-degree when computing the percentage of adopting in-neighbors. Table 15 shows that results are virtually unchanged. We find a strong positive relationship between the number (or share) of adopting neighbors and one's adoption choice, with magnitudes almost identical to the main specification. Fifth, we test whether results are sensitive to the type of network used. In Table 16 we re-estimate our main specification (Table 2 in the main text) but instead of the union network, we consider each of the four types of network ties we use to construct that network.

	Dependent variable: adopt					
	t = 1	t = 3	t = 5	t = 1	t = 3	t = 5
	(1)	(2)	(3)	(4)	(5)	(6)
# adopting neighbors (β_1)	0.017**	0.001	−0.009**	0.004	−0.007	−0.014***
	(0.007)	(0.008)	(0.004)	(0.006)	(0.008)	(0.004)
# adopting neighbors × high-uptake (β_2)	0.021***	0.028**	0.035***	0.025***	0.031***	0.038***
	(0.006)	(0.011)	(0.008)	(0.006)	(0.010)	(0.007)
degree	0.002***	0.001***	0.001**	0.001**	0.001	0.0004
	(0.001)	(0.001)	(0.0004)	(0.001)	(0.001)	(0.0004)
$\beta_1 + \beta_2$	0.038***	0.029***	0.026**	0.029***	0.024***	0.024**
Controls	—	—	—	✓	✓	✓
Observations	3,019	3,019	3,019	3,019	3,019	3,019
R^2	0.141	0.082	0.066	0.278	0.187	0.145

Note: *p<0.1; **p<0.05; ***p<0.01

Table 12: **Main specifications, varying threshold for adoption.** Models (1) and (4) reproduce, respectively, models (1) and (2) from Table 2 in the main text. They use a threshold of t = 1+ use of the platform in the past 12 months to define an adopter. Other models use a threshold of t = 3, 5+ uses of the platform in the past 12 months to define an adopter. Results are robust to using different thresholds for adoption.

	Dependent variable: adopt			
	Parsimonious	Baseline	Parsimonious	Baseline
	(1)	(2)	(3)	(4)
# adopting neighbors (β_1)	0.705***	0.195		
	(0.190)	(0.188)		
# adopting neighbors × high-uptake (β_2)	−0.259	0.138		
	(0.164)	(0.179)		
% adopting neighbors (β_1)			1.923	−1.184
			(2.839)	(4.247)
% adopting neighbors × high-uptake (β_2)			4.604	5.052
			(3.109)	(4.509)
degree	0.016	0.010	0.047***	0.029***
	(0.010)	(0.010)	(0.008)	(0.007)
$\beta_1 + \beta_2$	0.446***	0.333***	6.527***	3.868**
Controls	—	✓	—	✓
Observations	3,019	3,019	3,019	3,019
Akaike Inf. Crit.	903.416	628.667	925.181	638.397
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 13: **Main specifications, logistic regression.** This table reproduces Table 2 in the main text but uses a logistic regression instead of a linear probability model. Coefficients are log-odds ratios. While we cannot reject that peer effects are significantly higher in high-uptake villages than in low-uptake villages, peer effects in high-uptake villages are significantly different from zero in high-uptake villages, and are not in low-uptake villages.

	Dependent variable: adopt			
	Parsimonious	Baseline	Parsimonious	Baseline
	(1)	(2)	(3)	(4)
# adopting neighbors (β_1)	0.017**	0.005		
	(0.007)	(0.006)		
# adopting neighbors × high-uptake (β_2)	0.022***	0.025***		
	(0.006)	(0.006)		
% adopting neighbors (β_1)			0.102*	0.034
			(0.060)	(0.058)
% adopting neighbors × high-uptake (β_2)			0.351***	0.265***
			(0.102)	(0.083)
degree	0.002***	0.001**	0.004***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
$\beta_1 + \beta_2$	0.039***	0.03***	0.453***	0.298***
Controls	—	✓	—	✓
Observations	2,991	2,991	2,991	2,991
R^2	0.144	0.279	0.120	0.263
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 14: **Main specifications, village B excluded.** This table reproduces Table 2 in the main text but excludes village B from the sample. Results are qualitatively similar.

	Dependent variable: adopt			
	Parsimonious	Baseline	Parsimonious	Baseline
	(1)	(2)	(3)	(4)
# adopting in-neighbors (β_1)	0.039***	0.025***		
	(0.009)	(0.007)		
# adopting in-neighbors × high-uptake (β_2)	0.010	0.014*		
	(0.009)	(0.008)		
% adopting in-neighbors (β_1)			0.107**	0.085
			(0.053)	(0.086)
% adopting in-neighbors × high-uptake (β_2)			0.115	0.078
			(0.072)	(0.094)
degree	0.001*	0.001	0.004***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
$\beta_1 + \beta_2$	0.048***	0.04***	0.222***	0.163**
Controls	—	✓	—	✓
Observations	3,019	3,019	2,824	2,824
R^2	0.135	0.281	0.107	0.266
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 15: **Main specifications, directed ties.** This table reproduces Table 2 in the main text but uses the directed union network instead of the undirected union network. Results are qualitatively similar.

	Dependent variable: adopt							
	Family	Family	Friends	Friends	Lender	Lender	Solver	Solver
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# adopting neighbors (β_1)	0.027**	0.012	0.029**	0.002	0.012	−0.0003	0.025***	0.007
	(0.012)	(0.008)	(0.015)	(0.012)	(0.018)	(0.016)	(0.008)	(0.006)
# adopting neighbors × high-uptake (β_2)	0.016	0.021*	0.039**	0.042***	0.041*	0.035*	0.013	0.020
	(0.015)	(0.012)	(0.017)	(0.014)	(0.023)	(0.019)	(0.013)	(0.013)
degree	0.006***	0.004***	0.008***	0.005***	0.011***	0.007***	0.003***	0.002*
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Controls	—	✓	—	✓	—	✓	—	✓
$\beta_1 + \beta_2$	0.043***	0.032***	0.068***	0.045***	0.053**	0.035**	0.039**	0.027**
Observations	3,019	3,019	3,019	3,019	3,019	3,019	3,019	3,019
R^2	0.082	0.254	0.096	0.257	0.109	0.265	0.104	0.259

Note: *p<0.1; **p<0.05; ***p<0.01

Table 16: **Main specifications, by network type.** This table reproduces Table 2 in the main text but considers undirected ties from each of our four types of networks. Results are qualitatively similar.

5.2 Causally identifying peer effects

5.2.1 Instrumental variable

Initial encouragements to adopt the technology might be endogenous. Patterns of social influence may be confounded by other effects: two peers may adopt the technology due to similar unobservable characteristics, or because they have been exposed to related unobservable shocks. We address this issue using a generalization of An (2016) instrumental variable (IV) approach. Consider the following linear model of adoption:

$$y = X\beta + \lambda My + \epsilon \tag{2}$$

with y a vector of outcomes of length N, with $y_i = 1$ if i adopts the platform, and 0 otherwise, X an $N \times K$ matrix of individual covariates, M an $N \times N$ adjacency matrix, with diagonal entries set to 0, and ϵ an error term. Formally, the problem is that the autoregressive term My is quite possibly correlated with the error term ϵ. To address the issue of endogeneity, An (2016) recommends using an instrument z that is correlated with y, but not with ϵ. Using the two stages least squares (2SLS) procedure, we estimate the following models with OLS:

$$y = X\beta_0 + \lambda_0 z + \epsilon_0$$
$$y = X\beta + \lambda z + \gamma M\hat{y} + \epsilon,$$

with $\hat{y} = X\hat{\beta}_0 + \hat{\lambda}_0 z$.[2]

Our instrument is the distance from one's household to the location of the venue that GAPP used to hold its U-Bridge inception meeting. The idea is that the shorter the distance to the meeting venue, the more likely a villager is to adopt U-Bridge, simply by increasing the likelihood that she attends the meeting and learns about the new political communication technology. For the instrument to be valid, the exclusion restriction must be satisfied; i.e., we must assume that j's distance to the location of GAPP's inception meeting does not affect i's adoption via alternative channels than j's influence on i. This would be the case if contacts tended to cluster around locations that were more or less exposed to the meeting. Encouragingly, we find little (-.04) correlation between physical distance and having a social tie.

We also conduct several placebo tests to further explore potential violations of the exclusion restriction by conducting several placebo tests. If the exclusion restriction holds, mean peer distance from the meeting should affect one's adoption decision, but should not affect other theoretically meaningful predictors of adoption, such as political participation, leadership status, or phone ownership. Table 19 shows that this is indeed the case.

Note, furthermore, that our IV specification matches imperfectly our main specification (main text, Table 2). In particular, we omit our geographic control and meeting attendance. Indeed, because the instrument is the random choice of a meeting location, these controls are post-treatment covariates in this approach, and should therefore be excluded from the model.

The results of our IV models, reported in Table 18, confirms our basic adoption model. Again, results suggest a better model for for absolute threshold model (column 1) as compared to the fractional model (column 2). F-tests suggest, however, that our instrument is rather weak, with F statistics below 10. Note, furthermore, that our IV estimates are larger than comparable OLS estimates. We believe that using distance to the meeting as an instrument magnifies the effect of meeting attendance, because it compounds the effect of all neighbors attending the meeting, which is a very important predictor of adoption. Furthermore, given that our instrument is weak, results should be interpreted with care.

[2] Let $\hat{\theta}_{2SLS} = (\hat{\beta}, \hat{\lambda}, \hat{\gamma})$, $H = (X, z, Wy)$, and $\hat{H} = (X, z, W\hat{y})$. The variance covariance matrix writes $\mathbb{V}(\hat{\theta}_{2SLS}) = \hat{\sigma}^2(\hat{H}^\intercal H)^{-1}$ with $\hat{\sigma}^2 = e^\intercal e/N$ and $e = y - H\hat{\theta}_{2SLS}$.

	Dependent variable: adopt	
	Parsimonious IV	IV
	(1)	(2)
distance to meeting (km)	−0.007	−0.009
	(0.007)	(0.007)
distance to meeting (km) × high-uptake	0.003	−0.003
	(0.009)	(0.009)
degree	0.004***	0.003***
	(0.0003)	(0.0003)
Village FE	✓	✓
Controls	_	✓
Observations	2,832	2,832
R^2	0.092	0.193
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 17: **Instrumental variable approach, first stage.** OLS estimates of first stage of 2SLS models in table 18. Controls include all usual controls (see section 5.4 in the main text for details), except for meeting attendance and spatial influence. The instrument is weak.

	Dependent variable: adopt					
	Parsimonious IV	IV	OLS	Parsimonious IV	IV	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
# adopting neighbors (β_1)	0.016	0.008	0.004			
	(0.014)	(0.011)	(0.008)			
# adopting neighbors × high-uptake (β_2)	0.029**	0.040***	0.025***			
	(0.013)	(0.010)	(0.008)			
% adopting neighbors (β_1)				0.147	−0.062	0.031
				(0.151)	(0.117)	(0.069)
% adopting neighbors × high-uptake (β_2)				−0.080	0.259	0.137*
				(0.206)	(0.160)	(0.083)
degree	0.002***	0.001***	0.002***	0.004***	0.003***	0.003***
	(0.0005)	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
distance to meeting (km)	−0.007	−0.009	−0.009	−0.007	−0.009	−0.009
	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
distance to meeting (km) × high-uptake	0.006	0.002	0.002	0.003	−0.003	−0.002
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
F statistic	0.698	3.493**	_	0.698	3.493**	_
$\beta_1 + \beta_2$	0.044***	0.048***	0.03***	0.067	0.197*	0.168***
Village FE	✓	✓	✓	✓	✓	✓
Controls	_	✓	✓	_	✓	✓
Observations	2,832	2,832	2,832	2,832	2,832	2,832
R^2	0.105	0.212	0.216	0.093	0.194	0.197
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 18: **Instrumental variable estimates.** 2SLS estimates (models 1, 2, 4, 5) and corresponding OLS estimates (models 3, 6). We report F-statistics for instrument strength. Controls include all usual controls (see section 5.4 in the main text for details), except for meeting attendance and spatial influence. Although the instrument is weak, there is evidence of peer effects in high-uptake villages, but not in low-uptake villages (save for model 4).

	Dependent variable:			
	adopt	pol. participation	leader	phone
	(1)	(2)	(3)	(4)
mean peer distance to meeting (km)	-0.021^{**}	-0.015	-0.007	0.015
	(0.008)	(0.033)	(0.014)	(0.018)
age	0.0001	0.004^{***}	0.006^{***}	-0.003^{***}
	(0.0002)	(0.001)	(0.0004)	(0.0005)
female	-0.041^{***}	-0.231^{***}	-0.094^{***}	-0.120^{***}
	(0.011)	(0.018)	(0.016)	(0.013)
income	0.003	0.046^{***}	0.016^{***}	0.027^{***}
	(0.003)	(0.010)	(0.005)	(0.006)
secondary education	0.122^{***}	0.174^{***}	0.002	0.310^{***}
	(0.014)	(0.020)	(0.022)	(0.018)
pro-sociality	-0.058^{**}	0.042	-0.016	-0.014
	(0.025)	(0.053)	(0.023)	(0.037)
Constant	0.105^{***}	-0.260^{***}	0.022	0.217^{***}
	(0.024)	(0.088)	(0.039)	(0.057)
Observations	2,832	2,832	2,832	2,832
R^2	0.105	0.123	0.099	0.229
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 19: **Placebo tests.** OLS estimates with village-level fixed effects; standard errors in parentheses. Mean peer distance to meeting location affects adoption (model 1), but does not affect political participation (model 2), being a leader (model 3), or using a phone (model 4).

5.2.2 Degree and other network characteristics

As Aronow and Samii (2017) argue, exposure to peer influence is endogenous to one's network position. If the technology diffuses through patterns of social influence, then individuals with more central network positions are more likely to be exposed to such influence. In the limit, agents with no neighbors cannot be exposed to any influence, while agents with many neighbors are subjected to much influence. To address this issue, we compare between individuals with similar network positions. Yet, network positions can only be described partially, using centrality scores that each capture different aspects of one's network position.

	Dependent variable: adopt	
	OLS	*GAM (continuous)*
	degree strata	GAM
	(1)	(2)
# adopting neighbors (β_1)	0.008	0.003
	(0.005)	(0.006)
# adopting neighbors × high-uptake (β_2)	0.023***	0.026***
	(0.006)	(0.006)
degree $\in [8, 9]$	0.0003	
	(0.008)	
degree = 10	−0.003	
	(0.013)	
degre $\in [11, 12]$	−0.008	
	(0.011)	
degree = 13	−0.016**	
	(0.006)	
degre $\in [14, 15]$	−0.007	
	(0.007)	
degre $\in [16, 17]$	−0.016**	
	(0.007)	
degre $\in [18, 20]$	0.003	
	(0.016)	
degre $\in [21, 25]$	0.003	
	(0.015)	
degree > 25	0.056***	
	(0.021)	
$\beta_1 + \beta_2$	0.031***	0.029***
Controls	✓	✓
Observations	3,019	3,019
R^2	0.280	
UBRE		0.031
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 20: **Flexible controls for degree.** Model 1 controls for degree using degree strata based on sample deciles. Degree < 8 is the reference category. Model 2 uses a generalized additive model, and controls for degree using thin-plate regression splines. Our finding that peer effects are larger in high-uptake villages is robust to such controls.

We re-estimate our baseline specification but control very flexibly for one major centrality score: degree centrality. Table 20 reports the results from two specifications. One uses degree strata that split the population into degree deciles. The second controls from degree non-parametrically using generalized additive modeling, with thin-plate splines. Peer influence is robust to such controls.

Second, we re-estimate our baseline by controlling for a host of standard centrality scores: degree, betweenness, closeness, eigenvector, Bonacich centralities, and clustering. Eigenvector and Bonacich centralities are recursive metrics where a node is more central to the extent that it is connected to more central node. Other concepts are defined in section 1 of this SI. We estimate one model per centrality score, and divide the population in three strata based on which tercile they belong to. Again, peer influence is robust to such controls.

	Dependent variable: adopt					
	Degree	Betweenness	Closeness	Eigenvector	Bonacich	Clustering
	(1)	(2)	(3)	(4)	(5)	(6)
# adopting neighbors (β_1)	0.010**	0.011**	0.011**	0.009**	0.011**	0.011**
	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)
# adopting neighbors × high-uptake (β_2)	0.024***	0.024***	0.024***	0.024***	0.024***	0.023***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
medium centrality	−0.008	−0.017**	0.004	0.001	0.006	−0.019***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.010)	(0.005)
high centrality	0.009*	−0.002	−0.001	0.017	−0.00004	−0.017***
	(0.005)	(0.008)	(0.005)	(0.011)	(0.008)	(0.006)
Controls	✓	✓	✓	✓	✓	✓
$\beta_1 + \beta_2$	0.034***	0.035***	0.035***	0.034***	0.035***	0.034***
Observations	3,019	3,019	3,019	3,019	3,019	3,019
R^2	0.274	0.275	0.273	0.274	0.274	0.275
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 21: **Network covariates.** OLS estimates with village-level fixed effects. Network controls use tercile strata, with the lowest tercile as the reference category. Peer effects are robust to these network controls. Consistently with simple contagion, low clustering individuals adopt more. Our finding that peer effects are larger in high-uptake villages is robust to such controls.

5.2.3 Matching

Following Aral, Muchnik and Sundararajan (2009), we use matching to address simultaneously the problems of endogenous initial encouragements to adopt the technology, and endogenous exposure to peer influence due to network position. Our matched sample matches both on individual and network characteristics. We selected individual characteristics that are substantially and theoretically meaningful predictors of uptake: phone ownership, secondary education, political participation, and meeting attendance. Our network characteristics are degree and eigenvector centrality. Matching alleviates bias by constructing a treatment and a control group that more are comparable on such observable characteristics. It has the additional benefit of constructing groups that are also presumably more comparable on other network characteristics, since centrality scores tend to be highly correlated.

Matching requires using a binary treatment. As in Aral, Muchnik and Sundararajan (2009), we make our treatment binary by defining cutoffs in the number of adopting neighbors above which we consider that an observation is treated. Specifically, we use cutoffs of one, two, and three neighbors. Having defined these cutoffs, we compared, for each cutoff, three different matching procedures: neighbor, coarsened exact, and full matching. We chose full matching because it is the procedure that achieved the highest distance reduction for all three cutoffs. Figure 6 shows the results of our matching procedure for a cutoff of one neighbor. Our matched sample is balanced on all characteristics but political participation and belonging to a high-uptake village.

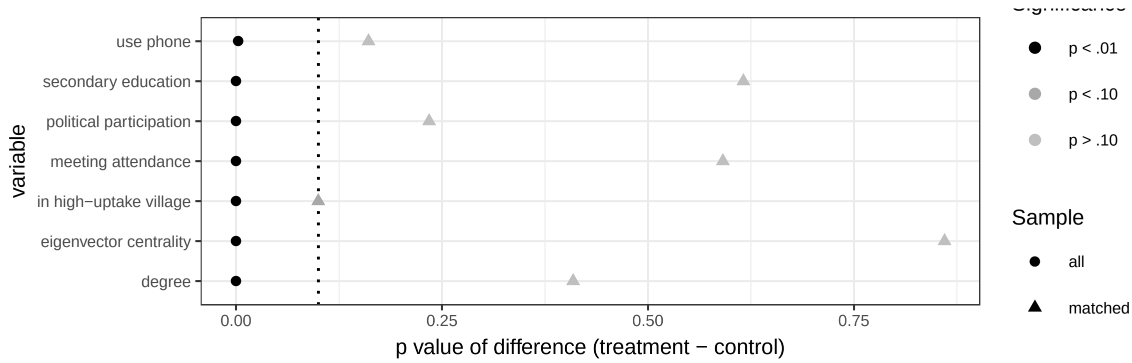

Figure 6: **Covariate balance on dimensions used for matching ($t = 1$).** We report the p-value of the difference in means in the treatment and control group in the full and matched sample, using full matching. Matching achieves balance in degree, phone use, and meeting attendance.

We then re-estimate our baseline specification on each matched sample. Table 22 shows that our results are largely robust to using a matched sample. The magnitude of the treatment effect increases with the strength of the treatment in high-uptake villages. It remains low and loses statistical significance in low-uptake villages.

	Dependent variable: adopt		
	Matched sample, $t = 1$	Matched sample, $t = 2$	Matched sample, $t = 3$
	(1)	(2)	(3)
# adopting neighbors $\geq t$ (β_1)	0.026**	−0.021	0.011
	(0.010)	(0.017)	(0.021)
# adopting neighbors $\geq t \times$ high-uptake (β_2)	0.023	0.094***	0.085***
	(0.020)	(0.022)	(0.018)
degree	0.004***	0.003***	0.002**
	(0.001)	(0.001)	(0.001)
$\beta_1 + \beta_2$	0.049**	0.073***	0.097***
Controls	✓	✓	✓
Observations	3,019	3,019	3,019
R^2	0.291	0.287	0.248
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 22: **Matching estimates.** OLS estimates with village-level fixed effects on a matched sample using full matching (see figure 6 for details on balance and section 5.4 in the main text for details about estimation). Treatment is an indicator variable that equals 1 if i has at least t adopting adopting neighbors. Our finding that peer effects are larger in high-uptake villages is robust using a matched sample.

6 Additional evidence on the empirical implications of the model

This section provides additional evidence supporting our conclusions regarding the empirical implications of the model (Section 6 in the paper).

6.1 Two-stage selection model

We also probe into the mechanism underlying neighbors' influence. Do neighbors foster adoption by spreading news about the existence of the technology, or by pushing individuals who already know of the innovation to adopt it? We answer this question by estimating a type-2 Tobit model for binary outcomes (Cameron and Trivedi, 2005). This selection model separates the fact of having heard about the platform from the decision to adopt it. Our outcomes are:

$$y_{1i} = \begin{cases} 1, & \text{if } i \text{ hears about the platform} \\ 0, & \text{otherwise} \end{cases}$$

$$y_{2i} = \begin{cases} 1, & \text{if } i \text{ adopts the platform and } y_{1i} = 1 \\ 0, & \text{if } i \text{ does not adopt the platform and } y_{1i} = 1 \\ \_, & \text{if } y_{1i} = 0 \end{cases}$$

That is, deciding whether to adopt the platform (y_{2i}) is defined if and only if one heard about it ($y_{1i} = 1$). Let x_{1i} and x_{2i} be column vectors of individual-level predictors for hearing about the platform and adopting it, respectively. We consider the following binary type-2 Tobit model:

$$\underbrace{p(y_{2i}|x_{1i}, x_{2i})}_{\text{adopting}} = \underbrace{p(y_{2i}|x_{2i}, y_{1i})}_{\text{adopting conditional on hearing}} \underbrace{p(y_{1i}|x_{1i})}_{\text{hearing}} \tag{3}$$

We estimate the model using logistic regression, and account for village-level effects by adding village indicators. This model can easily be estimated using two logistic regressions: the first regresses y_1 on X_1 for the whole sample, and the second regresses y_2 on X_2 for those observations where $y_{1i} = 1$. In the first stage, we regress hearing about the technology on the number of hearing neighbors. In the second stage, we regress adopting the technology on the number of adopting neighbors. We use the same set of controls as in the main text, with the exception that we exclude meeting attendance from the first stage, because it perfectly predicts hearing about the platform. We also estimate a reduced-form specification identical to our baseline specification using logistic regression (Table 2, model 2) to compare effect sizes in the second stage to a corresponding reduced-form specification. Table 23 reports the results. Neighbors influence both hearing about the technology but only in high-villages do they affect adopting it (as derived from our theoretical model). Both the reduced for and the selection model arrive at the same conclusion: while we cannot reject the null that peer effects are comparable in high- and low-uptake villages, we do find that peer effects in high-uptake villages are significantly different from zero, which they are not in low-uptake villages.

	Dependent variable:		
	heard	adopt	
	First stage	Second stage	Reduced form
	(1)	(2)	(3)
# hearing neighbors (β_1)	0.155^{***}		
	(0.033)		
# hearing neighbors × high-uptake (β_2)	0.0003		
	(0.031)		
# adopting neighbors (β_1)		0.139	0.195
		(0.152)	(0.188)
# adopting neighbors × high-uptake (β_2)		0.168	0.138
		(0.164)	(0.179)
degree	−0.0003	0.009	0.010
	(0.010)	(0.009)	(0.010)
Controls	✓	✓	✓
$\beta_1 + \beta_2$	0.156^{***}	0.307^{***}	0.333^{***}
Observations	3,019	938	3,019
Akaike Inf. Crit.	3,021.841	535.707	628.667
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 23: **Two-stage selection model.** Logistic regression estimates with village-level fixed effects. Coefficients are log-odds ratios. Model (1) is the first stage (hearing about the platform), model (2) is the second stage (adoption conditional on hearing), and model (3) is the corresponding reduced-form model (reproduced from Table 13 in the SI). While we cannot reject the null that peer effects are comparable in high- and low-uptake villages, peer effects in high-uptake villages are significantly different from zero, which they are not in low-uptake villages.

6.2 Empirical implications of the model

Table 24 provides evidence for signal discounting in low-uptake villages. We estimate separately the effect of peers that state being satisfied by the platform to those that do not. We find that in both high and low-uptake villages, adoption increases in the number of satisfied and unsatisfied neighbors, to a comparable magnitude (tests $H_0 : \beta_1 \neq \beta_3$ and $H_0 : \beta_1 + \beta_2 \neq \beta_3 + \beta_4$). These effects, however, vanish in low-uptake villages, where they are not significantly different from zero. In principle, one would test whether messages from dissatisfied neighbors tend to decrease adoption, and messages from satisfied neighbors to increase adoption, and whether the magnitude of these effects is significantly lower in low-uptake villages. However, our measure of satisfaction may be a serious underestimate, because it was collected two years after the start of the program. Satisfaction was probably higher then, as was the importance of communication among neighbors. As such, we interpret results with caution. Some evidence is consistent with the theory: both satisfied and non-satisfied neighbors exert influence in high-uptake villages, but not in low-uptake villages, while no evidence explicitly contradicts it.

	Dependent variable: adopt		
	Low-uptake	High-uptake	Full sample
	(1)	(2)	(3)
# non-satisfied adopting neighbors (β_1)	0.009	0.030***	0.006
	(0.009)	(0.005)	(0.007)
# non-satisfied adopting neighbors × high-uptake (β_2)			0.026***
			(0.009)
# satisfied adopting neighbors (β_3)	0.001	0.025**	−0.001
	(0.010)	(0.010)	(0.011)
# satisfied adopting neighbors × high-uptake (β_4)			0.027**
			(0.013)
degree	0.002*	0.001	0.001**
	(0.001)	(0.001)	(0.001)
Controls	✓	✓	✓
$H_0 : \beta_1 \neq \beta_3$	0.42	0.23	0.27
$H_0 : \beta_1 + \beta_2 \neq \beta_3 + \beta_4$			0.32
$\beta_1 + \beta_2$			0.031***
$\beta_3 + \beta_4$			0.026***
Observations	1,358	1,661	3,019
R^2	0.239	0.291	0.279
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 24: **Signal discounting in high- and low-uptake villages.** All models decompose the effect of peers into the effect of a peer adopting (contagion) and a peer adopting and declaring being satisfied (satisfaction). Models 1 and 2 are estimated separately for low- and high-uptake villages, while model 3 is estimated for the entire sample. The contagion and satisfaction effects are significantly larger in high-uptake villages than in low-uptake ones, lending support to our discounting hypothesis.

Table 25 shows finding for the model prediction that *strong ties* are more effective than weak ties in supporting truthful communication.

	Dependent variable: adopt	
	Simple vs. complex ties	Types of relationships
	(1)	(2)
# adopting simple ties, β_s	0.024***	
	(0.004)	
# adopting simple family		0.013
		(0.008)
# adopting simple friends		0.034**
		(0.014)
# adopting simple lender		0.013
		(0.013)
# adopting simple solver		0.011
		(0.009)
# adopting complex ties, β_c	0.027***	0.013***
	(0.006)	(0.004)
degree	0.001**	0.001**
	(0.001)	(0.001)
$\beta_c - \beta_s \neq 0$, F statistic	0.522	—
Controls	✓	✓
Observations	3,019	3,019
R^2	0.275	0.278
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 25: **Strength of network ties.** Model 1 reports the effect of one adopting neighbor with whom i shares a single type of relationship (simple tie) and more than one type of relationship (complex tie). Models 2 breaks down simple ties into family, friends, leader, and solver. Complex ties are weakly more influential than simple ties (F-test not significant). Friendship and family ties are the most influential simple ties.

6.2.1 Informal institutions and peer effects

In this section we further investigate the mediator between village-level peer effects and village characteristics. The paper discusses the following factors:

- Ethnic concentration
- Religious concentration
- Leadership concentration. To compute this measure, we use results from a public goods game where treated respondents could designate a person in the village that would receive all the donations. We then compute the concentration index in the number of mentions for each leader (i.e. nominated individual). We compute several versions of this metric using thresholds in the number of mentions above which we consider a nominated individual to be a leader. In the baseline ($t = 1$), we define an individual to be a leader if she has been mentioned at least once. We then define an individual to be a leader if she has been mentioned at least twice, thrice, or four times ($t = 2, 3, 4$ respectively).
- Pro-sociality. We compute this measure using the mean donation at the village level in a public goods game and in a dictator's game.

Figure 7 reports the distribution of such mediators separately for high- and low-uptake villages. We find that high and low-uptake villages differ in mostly in their leadership concentration, but do not differ in their other mediators. Tables 26 and 27 report the effect of such mediators on the outcome.

First, we find that irrespective of how we define leadership concentration, it has a positive and significant (interactive) effect on the size of peer effects (Table 26). Second, we find (Table 27, columns 3-4) that pro-sociality has a large, positive and significant effect on the size

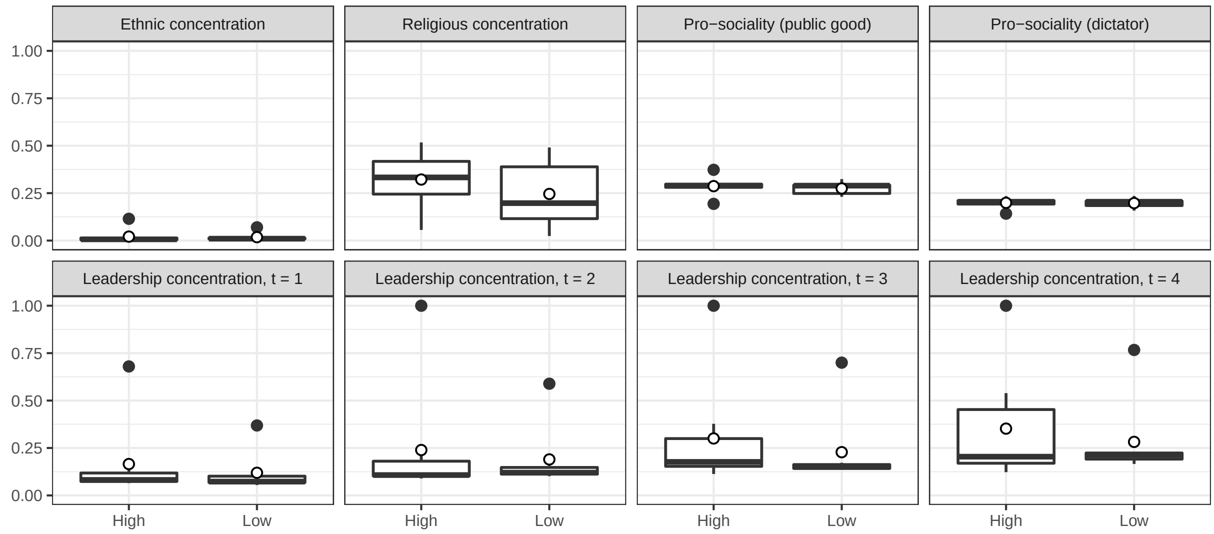

Figure 7: **Mediators in high and low uptake villages.** White points represent the mean. The distribution of mediators is comparable in high and low uptake villages for ethnic and religious concentration and for pro-sociality. High-uptake villages have more concentrated leadership than low-uptake villages. The difference is increasingly strong for higher thresholds.

	Threshold:			
	$t = 1$	$t = 2$	$t = 3$	$t = 4$
	(1)	(2)	(3)	(4)
leadership concentration × # adopting neighbors	0.077	0.051	0.050	0.045
	(0.027, 0.126)	(0.017, 0.084)	(0.021, 0.078)	(0.019, 0.070)
leadership concentration	−0.045	−0.034	−0.019	0.008
	(−0.235, 0.207)	(−0.161, 0.128)	(−0.148, 0.152)	(−0.126, 0.183)
degree	0.001	0.001	0.001	0.001
	(0.001, 0.002)	(0.001, 0.002)	(0.001, 0.002)	(0.001, 0.002)
Controls	✓	✓	✓	✓
Random intercept	✓	✓	✓	✓
Random slope	✓	✓	✓	✓
Observations	2,991	2,991	2,991	2,991

Note: 95 percent credible intervals in parenthesis.

Table 26: **Random effect models, leadership concentration.** All models include a random slope for the number of adopting neighbors by village. These models investigate how much the peer effect on hearing and adopting differ as a function of leadership concentration using an interaction term. Each column uses a different threshold for computing leadership concentration. In all specifications, higher leadership concentration leads to higher peer effects.

of peer effects. This is obviously consistent with our model. However, we also find (Figure 7) that while leadership concentration is larger in high-uptake villages, such villages do *not* exhibit greater level of pro-sociality as compared to low-uptake villages. Together these findings suggest that unlike leader concentration, pro-sociality cannot account for difference between low and high-uptake villages.

	Dependent variable: adopt			
	Religion	Ethnicity	Pro-sociality (1)	Pro-sociality (2)
	(1)	(2)	(3)	(4)
religious concentration × # adopting neighbors	0.047			
	(0.006, 0.085)			
religious concentration	0.027			
	(−0.175, 0.263)			
ethnic concentration × # adopting neighbors		0.147		
		(−0.223, 0.523)		
ethnic concentration		1.187		
		(0.379, 2.408)		
pro-sociality (public goods) × # adopting neighbors			0.062	
			(0.019, 0.100)	
pro-sociality (public goods)			0.666	
			(−0.012, 1.565)	
pro-sociality (dictator) × # adopting neighbors				0.089
				(0.027, 0.145)
pro-sociality (dictator)				0.758
				(−0.399, 2.238)
degree	0.001	0.002	0.001	0.001
	(0.001, 0.002)	(0.001, 0.002)	(0.001, 0.002)	(0.001, 0.002)
Controls	✓	✓	✓	✓
Random intercept	✓	✓	✓	✓
Random slope	✓	✓	✓	✓
Observations	2,991	2,991	2,991	2,991
Note:				95 percent credible intervals in parenthesis.

Table 27: **Random effect models, other mediators.** All models include a random slope for the number of adopting neighbors by village. Do peer effect on hearing and adopting differ as a function of the value of mediators? Higher values of pro-sociality and religious concentration induce larger peer effects, while higher ethnic concentration does not.

References

An, Weihua. 2016. "On the Directionality Test of Peer Effects in Social Networks." *Sociological Methods & Research* 45(4):635–650.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103(484):1481–1495.

Aral, Sinan, Lev Muchnik and Arun Sundararajan. 2009. "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks." *Proceedings of the National Academy of Sciences* 106(51):21544–21549.

Aronow, Peter M and Cyrus Samii. 2017. "Estimating Average Causal Effects under General Interference, with Application to a Social Network Experiment." *Annals of Applied Statistics* 11(4):1912–1947.

Bala, Venkatesh and Sanjeev Goyal. 1998. "Learning from neighbours." *The review of economic studies* 65(3):595–621.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo and Matthew O Jackson. 2013. "The Diffusion of Microfinance." *Science* 341(6144):1236498.

Blackwell, David et al. 1951. Comparison of Experiments. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability.* The Regents of the University of California.

Cameron, Adrian Colin and Pravin K. Trivedi. 2005. *Microeconometrics - Methods and Applications.* Cambridge University Press.

DeGroot, Morris H. 1974. "Reaching a consensus." *Journal of the American Statistical Association* 69(345):118–121.

Larson, Jennifer M, Janet I Lewis and Pedro Rodrıguez. 2017. "From Chatter to Action: How Social Networks Inform and Motivate in Rural Uganda." *Working paper* .
URL: *http://bit.ly/2DwT2Lt*