AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

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

**Romain Ferrali** New York University Abu Dhabi
**Guy Grossman** University of Pennsylvania
**Melina R. Platas** New York University Abu Dhabi
**Jonathan Rodden** Stanford University

**Abstract:** *Do social networks matter for the adoption of new forms of political participation? We develop a formal model showing that the quality of communication that takes place in social networks is central to understanding whether a community will adopt forms of political participation where benefits are uncertain and where there are positive externalities associated with participation. Early adopters may exaggerate benefits, leading others to discount information about the technology's value. Thus, peer effects are likely to emerge only when informal institutions support truthful communication. We collect social network data for 16 Ugandan villages where an innovative mobile-based reporting platform was introduced. Consistent with our model, we find variation across villages in the extent of peer effects on technology adoption, as well as evidence supporting additional observable implications. Impediments to social diffusion may help explain the varied uptake of new and increasingly common political communication technologies around the world.*

**Verification Materials:** The data and materials required to verify the computational reproducibility of the results, procedures and analyses in this article are available on the *American Journal of Political Science* Dataverse within the Harvard Dataverse Network, at: https://doi.org/10.7910/DVN/NOYBCQ.

Political participation is costly, and benefits of participating are often uncertain. If I participate in a protest, will it lead to a policy change? If I vote, will it lead to a change of government? If I report a problem about a public school, will the problem be solved? All of these types of political activities are characterized by an additional core feature: positive externalities. My political action may be welfare-improving not only for me, but also for others, and returns from participation depend on the actions of other agents. The decision about whether or not to take a costly political action under uncertainty thus hinges not only on what I expect others to do (coordination), but also on the information I gather about the expected benefits (communication). Acquiring information about potential benefits is particularly important for new forms of political participation—voting for the first time in a newly democratic state, contacting political leaders on social media, or sending text messages to report potholes or broken streetlights.

In this article, we develop and empirically test a model that brings together insights from hitherto distinct literatures on political participation and technology adoption to explain community- and individual-level variation in new forms of political engagement. The key insight that

Romain Ferrali is Postdoctoral Associate, New York University Abu Dhabi, Division of Social Science, Building A5-1135, Abu Dhabi, United Arab Emirates (rferrali@nyu.edu). Guy Grossman is Associate Professor, Department of Political Science, University of Pennsylvania, 133 S. 36th Street, Philadelphia, PA 19104-6215 and EGAP member (ggros@sas.upenn.edu). Melina R. Platas is Assistant Professor, New York University Abu Dhabi, Division of Social Science, Building A5-145, Abu Dhabi, United Arab Emirates; and EGAP member (mplatas@nyu.edu). Jonathan Rodden is Professor, Stanford University, Department of Political Science, Encina Hall Central, 616 Serra Street, Stanford, CA 94305 (jrodden@stanford.edu).

We are thankful to The United States Agency for International Development, Social Impact and Stanford GSB for financial support. We gratefully acknowledge the cooperation of members of the Arua district government, as well as RTI, GAPP, UNICEF Uganda, USAID/Uganda, and the USAID Center of Excellence on Democracy, Human Rights, and Governance, without whom this study would not have been possible. We thank Innovations for Poverty Action Uganda for excellent data collection, and Jon Helfers, Maximillian Seunik, Areum Han, and Zachary Tausanovich for providing valuable research assistance at various stages of the project. We received helpful feedback from Daeyoung Jeong and Korhan Kocak, as well as seminar participants at Brigham Young University, the University of Pennsylvania, NYU Abu Dhabi, Princeton, Stanford, Emory, Essex, the University of Virginia, the London School of Economics, King's College London, and the Stockholm School of Economics.

***American Journal of Political Science***, Vol. 64, No. 3, July 2020, Pp. 536–553


DOI: 10.1111/ajps.12471

emerges from the model is that the *quality of communication* that takes place in social networks is central to understanding whether a community will adopt forms of political participation where benefits are highly uncertain and where there are positive externalities associated with participation. Our empirical analysis focuses on the adoption of a particular, but increasingly common, form of political participation: a new political communication technology (PCT) that allows users to report service delivery problems to their municipality or local government using digital/mobile technologies.

Our theory starts with the observation that positive externalities are the defining characteristic of many technologies of political participation—their benefit increases with the number of adopters. One person reporting a complaint is likely insufficient to induce a local government in a low-income country to address a service delivery problem like teacher absenteeism. One person with a picket sign is unlikely to change policy.

With *new* political technologies such as PCTs, citizens must also learn about the costs and benefits of that technology for widespread adoption to occur. This learning happens by sharing information on social networks. But unlike the adoption of widely studied private goods, such as new agricultural practices, with PCTs, citizens must believe not only that the technology is sound, but also that many others will adopt it.

We further argue that widespread adoption of such technologies depends on features of the relevant social network—namely, its ability to facilitate truthful communication. Given positive externalities, early adopters of a new PCT have incentives to exaggerate the benefits of adoption in order to encourage others to adopt. Recognizing this incentive, citizens may discount information they receive from early adopting peers. Not all social networks overcome this challenge of truthful communication, and if they do not, social diffusion does not occur.

In other words, peers help diffuse new technologies for political engagement that are defined by large positive externalities, but *only in networks where truthful communication is supported*, for example, by formal and informal institutions. In essence, we argue that the social diffusion process that underlies the adoption of new technologies is governed by an interaction between the nature of the technology (its associated externalities) and the characteristics of the network (its ability to support truthful communication).

Our model applies to a broad class of political participation defined by three features: costly behavior, uncertainty over the benefits of participation, and the presence of positive externalities. The particular type of political participation that motivates the theory and empirics of this study is the use of PCTs that are increasingly common in both developed and developing countries. From the British FixMyStreet platform (Sjoberg, Mellon, and Peixoto 2017) to text-messaging systems that rate public officials in Pakistan (Bhatti, Kusek, and Verheijen 2014), digital technologies allow for more frequent and cheaper forms of participation than traditional means of political engagement.

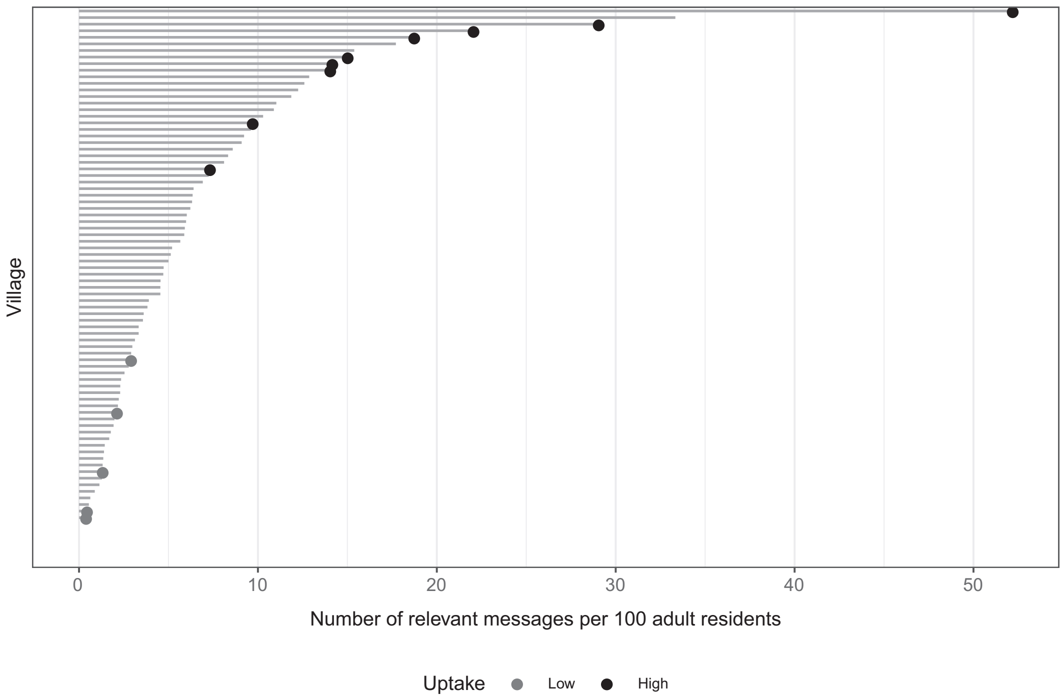
PCTs have the potential to transform the relationship between citizens and their governments, and to address some of the most intractable governance challenges. The potential benefits of PCTs are especially large when it comes to persistent, acute service delivery failures in low-income countries. New PCTs allow citizens to report problems in a way that is immediate, inexpensive, and potentially anonymous (Blair, Littman, and Paluck 2019). As we demonstrate, however, uptake of these technologies is often uneven across communities, and where uptake is low, it is unlikely to yield benefits to the public (Peixoto and Sifry 2017). The welfare losses resulting from low uptake can be significant, even in the short run, and especially so in developing country contexts. For example, failure to report contaminated water sources can lead to loss of life in the short term, even if adoption picks up in the longer term.

We test the implications of our theory with a case study drawing on original fieldwork in Uganda, where a new PCT was introduced. First, we describe the program, U-Bridge, which allows citizens in one Ugandan district to report service delivery problems to local government officials by sending free and anonymous text messages. Second, we not only show large variation in the adoption of this new PCT, but also demonstrate that existing theories have a hard time explaining the observed adoption patterns. Third, we present a new theory that is better positioned to explain when communities adopt new forms of political participation like PCTs. We then provide evidence in support of the observable implications of the theory using network data, survey data, and behavioral experiments. Next, we present our main result: when goods feature externalities, peer effects are not ubiquitous. Finally, we show support for more specific implications of the theory, and then conclude.

## The Setting

The PCT we study, U-Bridge, was implemented in Arua, a relatively poor district located in northwestern Uganda, through a collaboration between the local government, a local nongovernmental organization, USAID, and

**FIGURE 1 Number of Relevant Messages (Normalized) by Village**




UNICEF. Through U-Bridge, anyone could contact district officials by sending a text message to a short-code number. Messages sent through this platform were both *free* and *anonymous*, lowering the cognitive, monetary, and social costs for reporting service delivery problems. District officials in both technical and political positions were provided with tablets that enabled them to access and respond to incoming messages.

U-Bridge was implemented using a field experimental research design, encouraging usage in 131 randomly selected villages across Arua district, organized around 24 clusters. Residents in treatment villages were invited to attend periodic community meetings in a central location within clusters of four to five neighboring villages. In these meetings, attendees received information about national service delivery standards and were informed about ways to communicate with local officials. Public officials also provided attendees with an overview of government efforts in service delivery, especially in response to previous text messages. The first round of meetings was held in late 2014 as part of the launch of U-Bridge, and subsequent meetings were held quarterly.

Figure 1 shows the total number of (relevant) messages sent via the U-Bridge platform for each of the villages in our study area in the first 15 months after its launch, suggesting large variation in adoption rates. This variation is especially striking given that all villages are located in rural parts of the same district.

To explain variation in U-Bridge uptake across individuals and villages, we collected administrative and original survey data, which we conducted 2 years after the program launch. The in-person survey, which took place in April and May 2016, was administered to every available adult in 16 treatment villages[1] and included questions about respondents' demographics, social ties, perceptions of the quality of public goods and the capacity of their local government, and U-Bridge knowledge and usage. We surveyed 3,184 individuals, covering about 82% of the adults residing in the surveyed villages.[2]

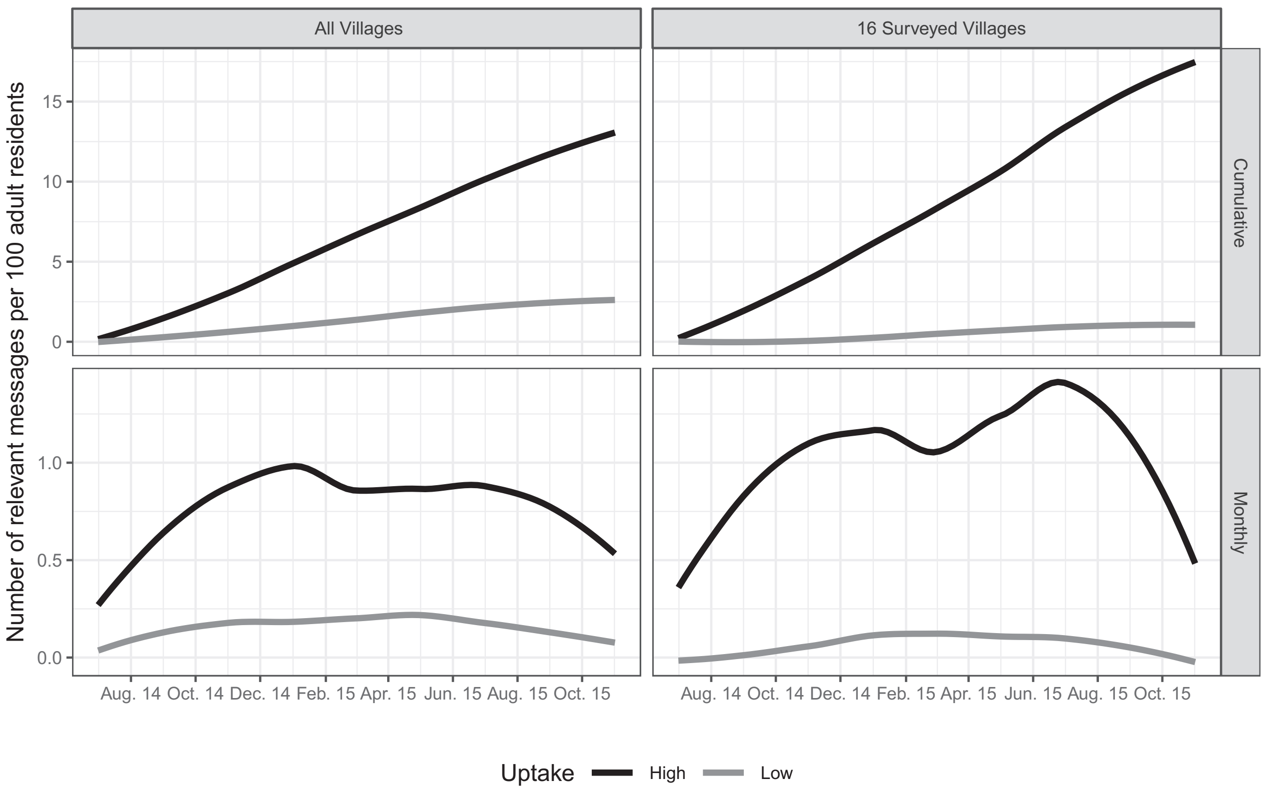
To maximize variation, about half (nine) of the study villages had a relatively high level of U-Bridge adoption (compared to what would be expected given village-level observable characteristics). The other half (seven) had relatively low adoption levels.[3] Figure 2

[1]The number of villages was determined by budget constraints.

[2]In Table 3 in the supporting information (SI), we report the number of individuals we surveyed in each village, the number of individuals mentioned by at least one person, and the number of adults living in each village, according to the 2014 census. This information allows calculating the number of missing respondents.

[3]To select villages, we regress the number of messages sent via U-Bridge (normalized by population) on village-level predictors and generate predicted values for the dependent variable ($\hat{y}$). We calculate the difference between the predicted value and the actual value of the dependent variable, that is, $\hat{\epsilon} = \hat{y} - y$, and use these residuals to select the highest- and lowest-performing villages

**FIGURE 2 Message Intensity Over Time**




*Note*: The monthly (bottom panels) and cumulative (top panels) number of relevant messages over time are shown per 100 residents, smoothed using a Loess fit. Villages in the left panel are clustered using a Gaussian mixture model with two mixture components (see SI Section 2.2 for additional details).

shows the cumulative number of relevant and actionable incoming messages between August 2014 and November 2015, broken down by village type (i.e., high or low uptake). The top panel shows the cumulative messages over time, whereas the lower panel shows messages sent by month. Messages in high-uptake villages increase for about 6 months, plateau, and then decline. By contrast, adoption in low-uptake villages never took off. What explains variation in adoption patterns?

## A Puzzle

We first test and reject several possible explanations for the patterns shown in Figure 2 by comparing high- and low-uptake villages across a variety of individual-level and village-level measures, as shown in Table 1.[4]

(largest positive and negative $\hat{\epsilon}$; SI Table 1). There are more high than low villages due to a replacement that took place during fieldwork.

[4]Additional details and measures available in SI Section 3.

1. **Heterogeneous demand:** We rule out the possibility that greater uptake of the PCT platform reflects greater demand for better public services. We find that, if anything, low-uptake villages should have greater demand for better services, as measured by the local goods index. We find no differences between high- and low-uptake villages with respect to the baseline quality of education services, a high-priority sector among message senders. We also find that residents across villages value similar services.
2. **Coordination failure:** We find no evidence of a coordination failure due to heterogeneous preferences in low-uptake villages. Within villages, villagers have high agreement on the types of public goods they value (SI Figure 4).
3. **Private vs. public goods:** Citizens may request personal favors when interacting with politicians. Perhaps villagers in high-uptake villages used U-Bridge to request private goods that have minimal externalities and are not subjected to

**Table 1 Descriptive Statistics**

| | Variable | Sample | High uptake | Low uptake | Δ | Std. diff. | min | max |
|---|---|---|---|---|---|---|---|---|
| **A. Individuals** | | | | | | | | |
| Outcome | % adopters | 0.044 | 0.063 | 0.020 | 0.043$^{***}$ | | 0 | 1 |
| | % heard | 0.306 | 0.348 | 0.249 | 0.100$^{*}$ | | 0 | 1 |
| | % satisfied | 0.392 | 0.439 | 0.207 | 0.232$^{**}$ | | 0 | 1 |
| Individual | age | 37.387 | 37.449 | 37.304 | 0.146 | | 18 | 101 |
| | % females | 0.579 | 0.564 | 0.596 | −0.032$^{**}$ | | 0 | 1 |
| | income | 2.550 | 2.636 | 2.436 | 0.201$^{**}$ | | 1 | 5 |
| | secondary education | 0.231 | 0.264 | 0.186 | 0.079$^{**}$ | | 0 | 1 |
| | % owns phone | 0.595 | 0.618 | 0.565 | 0.053 | | 0 | 1 |
| | % use phone | 0.174 | 0.192 | 0.15 | 0.043$^{*}$ | | 0 | 1 |
| | % leaders | 0.144 | 0.153 | 0.131 | 0.022 | | 0 | 1 |
| | political participation index | 0 | 0.060 | −0.079 | 0.139$^{**}$ | | −0.878 | 1.495 |
| | % attend meeting | 0.082 | 0.102 | 0.056 | 0.046$^{**}$ | | 0 | 1 |
| | pro-sociality | 0.200 | 0.202 | 0.198 | 0.004 | | 0 | 1 |
| Network | degree | 16.068 | 16.916 | 14.935 | 1.981$^{**}$ | | 1 | 227 |
| | betweenness | 0.007 | 0.006 | 0.007 | −0.001 | | 0 | 0.559 |
| | clustering coefficient | 0.386 | 0.385 | 0.388 | −0.003 | | 0 | 1 |
| Preferences | % education top priority | 0.356 | 0.377 | 0.328 | 0.049 | | 0 | 1 |
| *N* | | 3,184 | 1,820 | 1364 | 456 | | | |
| **B. Villages** | | | | | | | | |
| Design | % community meeting | 0.562 | 0.667 | 0.429 | 0.238 | 0.460 | 0 | 1 |
| | % dialogue meeting | 0.062 | 0.111 | 0 | 0.111 | 0.471 | 0 | 1 |
| Network | density | 0.098 | 0.114 | 0.078 | 0.036 | 0.462 | 0.052 | 0.405 |
| | path length | 2.117 | 2.091 | 2.151 | −0.060 | 0.376 | 1.602 | 2.327 |
| | global clustering | 0.251 | 0.266 | 0.230 | 0.036 | 0.463 | 0.174 | 0.549 |
| Population | adult population | 242.438 | 240.667 | 244.714 | −4.048 | 0.051 | 31 | 385 |
| | % employed | 0.862 | 0.855 | 0.870 | −0.015 | 0.170 | 0.679 | 1 |
| | % non-agriculture | 0.218 | 0.225 | 0.208 | 0.017 | 0.093 | 0 | 0.567 |
| | ethnic fractionalization | −0.069 | −0.057 | −0.086 | 0.029 | 0.105 | −0.480 | 0.468 |
| | poverty score | 0.044 | 0.064 | 0.018 | 0.045 | 0.475 | 0 | 0.406 |
| Politics | LC5 Chair turnout | 0.243 | 0.291 | 0.181 | 0.11$^{**}$ | 1.551$^{*}$ | 0.100 | 0.376 |
| | share LC5 winner | 0.679 | 0.614 | 0.763 | −0.149$^{**}$ | 1.091 | 0.303 | 0.883 |
| Distances | dist. to Arua (km) | 17.482 | 20.108 | 14.106 | 6.002 | 0.587 | 6.490 | 40.401 |
| | dist. to health center (km) | 1.042 | 0.929 | 1.189 | −0.260 | 0.264 | 0 | 3.692 |
| | dist. to school (km) | 0.865 | 1.141 | 0.512 | 0.629 | 0.687 | 0.031 | 3.651 |
| Public goods | public goods summary index | −0.002 | 0.093 | −0.123 | 0.215 | 0.394 | −1.280 | 1.232 |
| | local goods summary index | 0 | 0.284 | −0.366 | 0.650$^{**}$ | 1.63$^{*}$ | −0.615 | 1.241 |
| | teacher absenteeism | 0.722 | 0.723 | 0.722 | 0.001 | 0.007 | 0.525 | 1 |
| | students per class | 126.248 | 130.222 | 121.705 | 8.517 | 0.214 | 56 | 196.429 |
| *N* | | 16 | 9 | 7 | 2 | | | |

*Note*: The table reports mean values for the full sample, and for low- and high-uptake villages. Network characteristics are calculated from the union network. Difference in means are tested using a t-test, with standard errors clustered at the village level in panel A, and heteroskedastic robust standard errors in panel B. Due to small sample sizes, we also report standardized differences for village-level variables. A $*$ denotes that the 95% confidence interval around a standardized difference does not include 0.25. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

collective action problems. Coding each message according to the request "type," we find that the vast majority of messages sent via U-Bridge concerned substantive service provision problems, and very few were private requests (SI Figure 5).

4. **Government responsiveness:** Villagers are more likely to contact their local government if they expect greater responsiveness. This could be the case, for example, if clientelistic exchange took place at the community level (Rueda 2015) and high-uptake villages voted for the incumbent district chairperson at greater rates. Using 2016 election data, we find instead that incumbent vote share was somewhat *lower* in high-uptake villages. Thus, we find no evidence that high-uptake villages had reasons to expect greater government responsiveness.
5. **Different seeds:** Past work has highlighted the importance, for diffusion of information, of the identity and network position (Larson, Lewis, and Rodrıguez 2017) of initial "seeders." We compare the individual attributes and network characteristics of those attending GAPP's inception meetings and find small and insignificant differences in seeders' characteristics in high- and low-uptake villages (SI Table 9).
6. **Network properties:** Perhaps some networks do not facilitate processes of social diffusion due to "inadequate" structure. For example, Centola (2015) argues that diffusion processes are highly dependent on network properties—density, clustering, path length, and bridge width. When we examine core network-level properties (Table 1 above), we find substantively small differences between high- and low-uptake villages.

In developing a general theory for explaining variation in the adoption of new technologies, we use the above findings as our starting point. First, we assume homogeneous demand: high- and low-uptake villages have the same payoff function (everybody values the good equally—points 1 and 2). Second, we assume that taking action has positive externalities (people request public rather than private goods—point 3). We further assume that villagers face the same probability of having the public good delivered (government is equally responsive across villages—point 4). In addition, we assume that high- and low-uptake villages have similar "types" of early adopters (seeds' characteristics are similar—point 5) and similar network structures (point 6). Additional model assumptions are discussed below.

## Externalities, Networks, and Technology Adoption

Our model clarifies how externalities condition the role social networks play in technology adoption. In our model, agents decide at two time periods whether to adopt a new technology (or a good). Adoption is costly and yields benefits that depend on an unobserved state of the world that conditions how useful the technology is. Agents have heterogeneous prior beliefs about which state they are in. Agents are connected on a network and learn about the state of the world from previous waves of adoption, their personal experience with the good, and what their neighbors tell them about their experience. How adoption unfolds depends on the state of the world, prior beliefs about the technology, and, in the case of goods with externalities, whether a community is able to enforce truthful communication. In SI Section 4, we situate the model within the literature and prove the results described below.

### Setup

Consider a finite set of $N$ agents connected by the undirected graph $g = (G, N)$, where $G$ is a set of ties. There are three time periods $t \in \{0, 1, 2\}$ and an unobserved state of the world $\theta \in \{H, L\}$, drawn once at the beginning of the game. In the high state $H$, the technology is useful, whereas it is not in the low state $L$. In our context (a good with positive externalities), the high state means that a (local) government is both responsive to citizens' demands and capable of addressing them. The low state means that the government shows little responsiveness to those demands and/or lacks the capacity to address them. Each agent $i$ has prior over the state $\pi_i \equiv \Pr(\theta = H) \in (0, 1)$ and discounts the future with rate $\gamma \in (0, 1)$.

At time period $t = 0$, each agent $i$ may take the action $y_{i0} \in \{0, 1\}$. In our setting, *taking the action* ($y_{i0} = 1$) means sending a text message via the U-Bridge platform. The benefit $B_0 \in \{0, 1\}$ is then drawn with $\Pr(B_0 = 1|\theta) \equiv q(\theta, .) \in (0, 1)$. It is publicly observed, irrespective of one's adoption choice, and instantaneous payoffs accrue according to the payoff function $u(y_{i0}, \theta, .)$ that depends on $B_0$. As we detail below, the distribution of the benefit and payoffs crucially depend on whether the good has externalities. Agents who took the action get a private signal about the state $s_i \in \{0, 1\}$, with $\Pr(s_i = 1|\theta) = r_\theta \in (0, 1)$, representing private information early adopters get from their experience with the good. The private signal is informative: in the high (low) state, it is more likely than not to get a high (low) signal: $r_L < \frac{1}{2} < r_H$.

At time period $t = 1$, early adopters ($y_{i0} = 1$) simultaneously send messages $m_{ij} \in \{0, 1\}$ to their neighbors $j \in N_i(g)$ about their experience with the good. Agents may lie by sending some message $m_{ij} \neq s_i$, but they incur cost $\kappa \geq 0$ per lie. This parameter represents, in a reduced form, mechanisms that have been identified to sustain cooperation, such as moral costs of lying (Bénabou and Tirole 2011), third-party enforcement stemming from repeated interactions within communities (Fearon and Laitin 1996), or other behavioral mechanisms known to sustain punishment of defectors, such as inequality aversion (Engelmann and Strobel 2004).[5] Let $M_i^s \equiv \{m_{i'j} : i' = i\}$ be the (possibly empty) set of messages that $i$ sent. She gets payoff $v(M_i^s) = -\sum_{m \in M_i^s} 1\{m \neq s_i\}\kappa$, with $v(\emptyset) = 0$.

At time period $t = 2$, agents receive the (possibly empty) set of messages $M_i^r \equiv \{m_{i'j'} : j' = i\}$ and may again take the action $y_{i2} \in \{0, 1\}$. The benefit $B_2 \in \{0, 1\}$ is then drawn with the same distribution $q(\theta, .)$ as in $t = 0$ and is publicly observed. Payoffs then accrue according to the same payoff function $u(y_{i2}, \theta, .)$.

We now detail payoffs and the distribution of benefits in the cases with and without externalities. In both cases, adoption is costly and benefits depend on the state of the world. However, reaping the benefits further depends on the nature of the good. Without externalities, only adopters reap the benefit. With externalities, both adopters and non-adopters reap the benefit, but the probability of reaping such a benefit increases with the number of adopters.

***Goods without Externalities.*** Without externalities, payoffs only depend on one's actions: $q(\theta, .) = q(\theta)$, with $q(H) = p_H$ and $q(L) = p_L$ and $u(y_{it}, \theta, .) = u(y_{it}, \theta)$, with

$$u(y_{it}, \theta) = y_{it}(B_t - c). \tag{1}$$

Agents pay the cost of adoption $c$ and reap the benefit only if they adopt. We assume that the public signal conveyed by the benefit is informative, and that they have match-the-state utilities; that is, adoption is rational only in the high state. As such,

$$p_L < c < p_H;$$
$$p_L < \tfrac{1}{2} < p_H.$$

[5] At $t = 2$, agent $i$ is able to infer whether her neighbor $j$ lied to her by computing the likelihood of receiving a lie in the strategy profile under consideration given the rest of the information she accumulated, thus allowing for third-party punishment.

***Goods with Positive Externalities.*** With externalities, payoffs depend on the actions of other agents. Let $n_t \equiv \sum_{i \in N} y_{it}$ be the number of adopting agents in period $t$, and $n_{-it} \equiv n_t - y_{it}$ the number of adopting agents other than $i$ in period $t$. Then $q(\theta, .) = q(\theta, n_t)$ and $u(y_{it}, \theta, .) = u(y_{it}, \theta, n_{-it})$, with

$$u(y_{it}, \theta, n_{-it}) = B_t - y_{it}c. \tag{2}$$

Here, agents pay the cost of adoption only if they adopt, but they reap the benefit irrespective of their adoption choice. We assume again that the public signal conveyed by the benefit is informative and that agents have match-the-state utilities. Although the probability of reaping the benefit increases with the number of adoptions, this probability is lower in the low state. In our context, irrespective of the state of the world, the local government is more likely to deliver the benefit when receiving a large number of messages (Equation 3). For the same number of messages, however, the local government is less likely to deliver the benefit in the low state because it lacks capacity or will (Equation 4), to the point that sending any number of messages is too costly in the low state (Equation 5). Thus, for any $n \geq 0$:

$$q(\theta, n) < q(\theta, n + 1); \tag{3}$$

$$q(L, n) < \tfrac{1}{2} < q(H, n); \tag{4}$$

$$q(L, n) < c < q(H, n). \tag{5}$$

We simplify the problem by making a technical assumption—namely, that the marginal impact of an additional adopter in the high state on the probability of reaping the benefit is higher than in the low state: $q(H, n + 1) - q(H, n) > q(L, n + 1) - q(L, n)$.

Note that payoffs assume a constant adoption cost $c$ (Equations 1 and 2). This simplifying assumption encapsulates all differences among agents in their prior beliefs $\pi_i$. With externalities, this also simplifies interpretation by making more adoptions only increase the benefit, and not decrease costs.

In the case with externalities, there is no *a priori* reason to believe that messages are substitutes or that they are complements. Although Equation 3 requires that $q$ is strictly increasing in $n$, we do not make any assumption on its concavity or convexity. This accommodates cases where adoption decisions are complements (convexity), and cases where adoption decisions are substitutes, leading to collective action problems (concavity).

## Results

***Equilibrium.*** We now examine what drives adoption decisions both with and without externalities. In equilibrium, agents have threshold strategies: they adopt the technology if they are sufficiently certain to be in the high state. Consider an equilibrium profile $\sigma$. At each time period, agents choose the action that maximizes their expected payoff, using available information.

At $t = 0$, agents only rely on their prior. Early adopters are the agents who are sufficiently optimistic about the state; their prior $\pi_i$ exceeds some threshold $a^{\sigma}_{i0}$. How much optimism is required to trigger adoption depends on several factors. First, the threshold increases with the cost of adoption $c$. Second is the informativeness of the private signal. The more informative the signal (i.e., the higher its likelihood of matching the state), the lower the threshold. Indeed, if $i$ anticipates that she will get a very informative signal, she has an incentive to adopt early because that signal will allow her to discover the state more quickly. Third, agents consider the actions of other agents under profile $\sigma$. The threshold encapsulates whether adoption decisions are complements or substitutes,[6] and how much additional information she will obtain from her peers. For instance, should many agents adopt at $t = 0$ and truthfully communicate their signals to $i$, then adopting in the first stage would provide little additional information to $i$ for the second stage.

At $t = 2$, agents have more information and use it to inform their adoption decision. Let $S_{i2} \in \mathcal{I}_{i2}(y_{i0}, y_{-i0})$ be the vector of signals received by $i$ at $t = 2$. It contains $B_0$, the public signal received at $t = 0$; $M^r_i$, the vector of signals sent to $i$; and, if $y_{i0} = 1$, the private signal $s_i$. The set $\mathcal{I}_{i2}(y_{i0}, y_{-i0})$ is $i$'s *information structure* at time $t = 2$ and contains all potential realizations of $S_{i2}$, with $\mathcal{I}_{i2}(0, y_{-i0}) = \{0, 1\}^{|M^r_i|+1}$, and $\mathcal{I}_{i2}(1, y_{-i0}) = \mathcal{I}_{i2}(0, y_{-i0}) \times \{0, 1\}$. Agent $i$ adopts if her signals contain enough evidence favoring the high state, as captured by a higher (log) likelihood ratio under strategy profile $\sigma$, $l_{\sigma}(S_{i2}) \equiv \log[\frac{\Pr_{\sigma}(S_{i2}|\theta=H)}{\Pr_{\sigma}(S_{i2}|\theta=L)}]$. How much evidence is necessary depends on one's threshold $a^{\sigma}_{i2}(S_{i2})$. Similar to $t = 0$, agents who were originally too pessimistic about the state have higher thresholds, higher costs of adoption increase the threshold, and the threshold depends on the actions of other agents under profile $\sigma$. The following proposition encapsulates the discussion:

**Proposition 1 (Threshold Strategy).** *If strategy profile $\sigma$ is a perfect Bayesian equilibrium, then agents have a threshold strategy such that*

$$y^*_{i0} = 1 \iff \pi_i \geq a^{\sigma}_{i0} \text{ and}$$

$$y^*_{i2} = 1 \iff l_{\sigma}(S_{i2}) \geq a^{\sigma}_{i2}(S_{i2}),$$

*with $a^{\sigma}_{i2} : \mathcal{I}_{i2}(y^*_{i0}, y^*_{-i0}) \to \mathbb{R}$.*

***The Benefits of Truthful Communication.*** We now turn to the communication stage $t = 1$ and examine when agents may lie (e.g., misrepresent benefits). We define communication as *truthful* when all agents communicate information that matches their observed signal: $m_{ij} = s_i$. The value $V_{ig}$ of $i$'s information on graph $g$ is her expected payoff from all potential information she could receive $\mathcal{I}_{i2}(y_{i0}, y_{-i0})$ at $t = 2$, given that she responds optimally to that information. Formally, $V^{\sigma}_{ig}(y_{i0}, y_{-i0}) \equiv \sum_{S_{i2} \in \mathcal{I}_{i2}(y_{i0}, y_{-i0})} \mathbb{E}_{\theta}[u(y^*_{i2}(S_{i2}), \theta, .)|S_{i2}] \Pr_{\sigma}(S_{i2})$ is the *value of information structure* $\mathcal{I}_{i2}(y_{i0}, y_{-i0})$ under equilibrium profile $\sigma$ on graph $g$. In a perfect Bayesian equilibrium where communication is not truthful, agents misrepresent their signal with some probability. Intuitively, $i$'s information is most valuable under truthful communication because sharing inaccurate information introduces additional noise that makes inferences about the state less precise.[7] Formally:

**Proposition 2 (Truthful Communication Is Most Valuable).** *Consider equilibrium profile $\sigma_0$ with truthful communication and equilibrium profile $\sigma$ where some agent $j \in N_i(g)$ misrepresents her signal to $i$ with some probability. We have*

$$V^{\sigma}_{ig}(y_{i0}, y_{-i0}) \leq V^{\sigma_0}_{ig}(y_{i0}, y_{-i0}).$$

Without externalities, agents have no incentive to misrepresent because doing so brings no benefits. Agents are indifferent if there is no penalty for lying, but communication becomes truthful as soon as lying becomes costly:

**Proposition 3.** *Without externalities, truthful communication is a perfect Bayesian equilibrium for any $\kappa \geq 0$. It is the unique equilibrium for any $\kappa > 0$.*

With externalities, however, early adopters have an incentive to misrepresent their signal. Agent $i$ would like to benefit from the positive externality and, as such, gather as many adopters as possible in the second period, irrespective of whether she is set on adopting in the second period. In our context, even the early adopters that had poor private experiences with the technology have an

[6] If adoption decisions are complements, $i$'s threshold gets lower when she expects a larger number of adopters. If they are substitues, then $i$'s threshold gets higher when she expects a larger number of adopters.

[7] The claim only holds in the second stage. At $t = 0$, truthful communication might not be as valuable because the information that transits at $t = 1$ is a public good. Agents may delay adoption because they expect to benefit from alters' messages.

incentive to tell their neighbors that they had a good experience to push them to adopt in the second stage. If lying is punished with enough severity, that incentive disappears. However, the level of punishment required to restore truthful communication is generically higher for goods with externalities than for goods with no externalities:

**Proposition 4.** *With externalities, there are thresholds* $\bar{\kappa}_1, \bar{\kappa}_2$ *with* $0 \leq \bar{\kappa}_1 \leq \bar{\kappa}_2 \leq 1$ *such that truthful communication is a perfect Bayesian equilibrium if and only if* $\kappa \geq \bar{\kappa}_1$ *and is the unique perfect Bayesian equilibrium for any* $\kappa > \bar{\kappa}_2$.

Truthful communication has a key implication: It enables peer effects. Because neighbors share their experiences, they learn from the same sources of information and make more similar inferences. Such peer effects get stronger the more neighbors a dyad has in common, because the two neighbors acquire more similar information. Formally, this means that connecting two agents increases the correlation of their log-likelihood ratios:

**Proposition 5 (Truthful Communication Implies Peer Influence).** *Consider equilibrium profile* $\sigma_0$ *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* $j$. *Let* $\rho(x, y)$ *be the correlation coefficient between* $x$ *and* $y$ *and denote by* $S_{i2}^{g}$ *the set of signals received by* $i$ *at* $t = 2$ *on* $g$. *We have*

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

When communication is not truthful, agents put less weight on the messages sent by their neighbors when making inferences about state $\theta$. In the limit, the messages they receive are uninformative, and agents only use the public signal and their own private signal (if any) to derive the posterior. In this case, Proposition 5 no longer holds: The posteriors of neighbors are no more correlated than the posteriors of agents who are not connected on the social network.

## Informal Discussion of the Model

Based on our model, we chart several potential patterns of adoption over time, summarized in Figure 3. The *initial* adoption decision is driven by agents' priors, since they lack hard evidence at this stage (Proposition 1). If a village has many optimistic agents—agents who are sufficiently confident the government will be both *responsive* and *capable* of meeting their demands—there are many early adopters (top quadrants of Figure 3). Conversely, if a village has many pessimistic agents, then there are few early adopters (Figure 3, bottom quadrants).

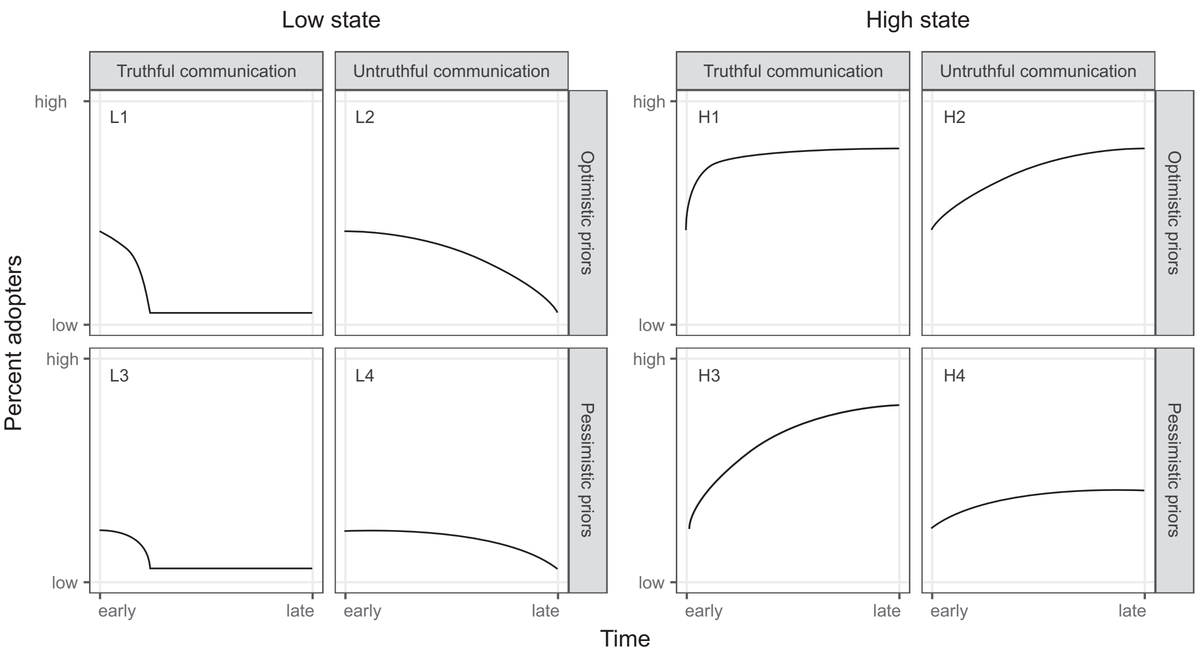
How adoption unfolds in later stages is a function of the state of the world because here, adoption depends on a richer informational environment (Proposition 1). This information is either gathered directly through public signals (e.g., whether the government offered adequate responses to problems raised by citizens), directly through private signals (one's own experience of using the platform), or indirectly through communication with peers. This information leads to convergence on the correct decision: in the high state, a series of good news leads villagers to adopt at high rates (Figure 3, right quadrants); in the low state, cumulative bad news leads villagers to low rates of adoption (Figure 3, left quadrants).

Importantly, truthful communication acts as a social multiplier; when agents believe each other, the information they exchange enables peer effects to kick in (Proposition 5). This, in turn, means that behavior adjusts faster to the state (Proposition 2). In Figure 3, quadrants with truthful communication (L1, H1, L3, and H3) adjust faster than, respectively, quadrants without truthful communication (L2, H2, L4, and H4). Untruthful communication leads to short-run welfare losses because agents reach the optimal outcome more slowly. The problem is particularly acute when the state is high and agents are pessimistic (H4), because it may result in long-run welfare losses. Pessimism prevents widespread initial adoption, and lack of truthful communication prevents the spread of information from early to late adopters. As such, agents may stick to their pessimistic priors and never adopt a technology that would have been beneficial.

Yet, unlike goods with no externalities, truthful communication may not always emerge when goods feature externalities (Propositions 3 and 4). As shown above, truthful communication only emerges when misrepresenting information is sufficiently costly. Together, our model yields several testable implications:

1. **Peer effects variability:** If there is truthful communication, then there are peer effects. If there is no truthful communication, then there are no peer effects.
2. **Discounting:** If there is no truthful communication, then agents discount peers' recommendations. If there is truthful communication, they do not discount peers' recommendations.
3. **Enforcement:** If there is a high cost of lying, then there should be truthful communication and peer effects. If there is a low cost of lying, then there should be no truthful communication and no peer effects.

**FIGURE 3 Illustration of Main Model Propositions**




*Note*: Initial adoption depends on priors (Proposition 1). Truthful communication enables peer effects (Proposition 5) that allow matching the state faster (Proposition 2).

4. **Initial adoption:** If agents have low priors, then initial adoption is low. If agents have high priors, then initial adoption is high.
5. **Convergence:** If there is truthful communication and given enough time, agents converge to the decision that matches the state of the world.

## Peer Effects Variability

To test the assumptions of the model and its main empirical implications, we use administrative data collected from Arua district local government, survey data from 16 villages where the new PCT platform was introduced, and focus group discussions (FGDs) with users and district officials. This section provides evidence supporting the model's assumptions and shows support for the broadest empirical expectation derived from the model: variation in peer effects, and hence uptake of the new technology, across village types.

### Model Assumptions

The validity of our model crucially depends on two core assumptions: (1) that sending messages through the U-Bridge platform is costly, and (2) that more messages being sent were expected to translate to a higher likelihood of government response. Qualitative evidence from FGDs with U-Bridge users suggests that these assumptions are met. For example, a major cost reported by villagers was the possibility that their identify as message senders would be revealed. Specifically, villagers expressed fear of retribution from the district government or street-level bureaucrats if their identities were known. One villager explained:

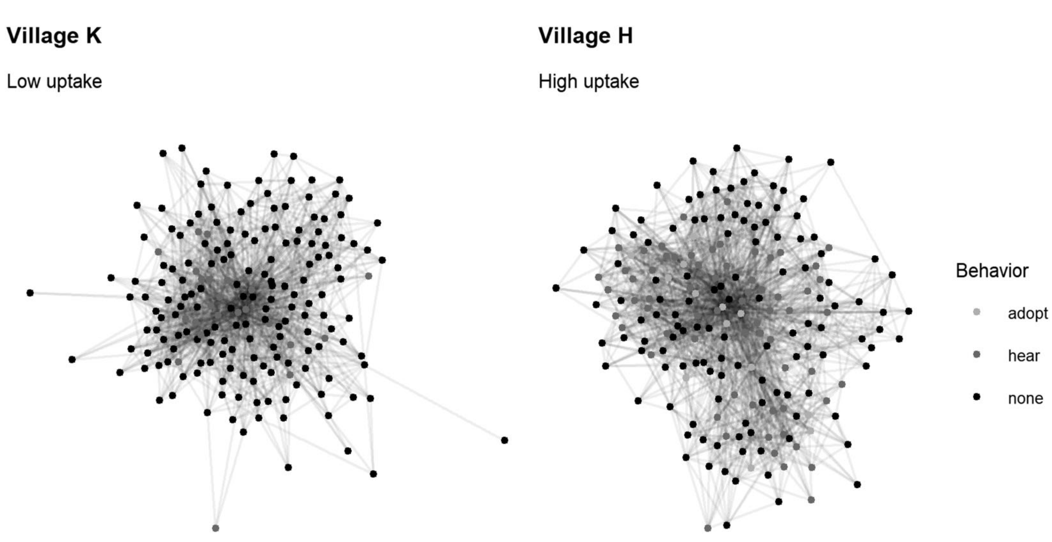
> If [our identities] are known, it would cause enmity between us since we are reporting mostly negative issues that might concern other people who have failed to do their jobs. [A lack of anonymity] would make us not send these messages.

As for our model's second assumption, many U-Bridge users we interviewed communicated clearly their belief that *collectively* sending messages was necessary for the program to succeed. As one user explained:

> I expected the government would respond because they said responses would be given after collecting many messages. So, if many people send the same message, then the district leaders will take action.

We now turn to an examination of cross-village variation in peer effects.

**FIGURE 4 Union Network of One High-Uptake Village (H) and One Low-Uptake Village (K)**




## Network Construction

We measure social networks using a standard name generator (Kolaczyk 2009) for four kinds of relationships: (1) *family* ties; (2) *friendship* ties; (3) *lenders*, to whom they would go to borrow money; and (4) *problem solvers*, to whom they would go to solve a problem regarding public services in the village. For each relationship type, respondents named up to five co-villagers. Note that some villagers (about 30%) were named by other respondents but not interviewed. Following standard practice (e.g., Larson and Lewis 2017), we exclude those nodes from the analysis.

We construct four "undirected" village networks for the four different types of ties, by collapsing directed ties into undirected ones. We further construct the union of those networks by defining a tie between $i$ and $j$ if there is at least one tie between them in any of the four networks. Figure 4 provides a graphical representation of the union network of two villages in the study area. Respondents who were knowledgeable about U-Bridge were asked to name the individuals from whom they heard about the platform. This allows tracking the diffusion process of knowledge about the new political communication system.

## Variable Description

Our key outcome measure is the adoption of U-Bridge. *Adopt* is a self-reported, binary variable that equals 1 if the respondent has used the platform at least once in the past 12 months. Similarly, *hear* is an indicator that gets the value of 1 if the respondent has heard about the U-Bridge service. By definition, U-Bridge adopters have a positive value for *hear*, but not vice versa. For those reporting that they have contacted the Arua district local government via U-Bridge (i.e., "adopters"), we also measure *satisfaction*: a binary variable that equals 1 if the respondent is at least somewhat satisfied with the platform.

Our key explanatory variables are network characteristics that support diffusion. We focus on two classes of diffusion models: (a) *fractional* threshold model, where an individual adopts a technological innovation if more than some *share* of her neighbors have adopted it (e.g., Acemoglu, Ozdaglar, and Yildiz 2011); and (b) *absolute* threshold model, where an individual adopts if more than some *number* of her neighbors have adopted (e.g., Centola and Macy 2007). When examining *absolute* contagion processes, our key independent variable, *# adopting neighbors*, counts, for each individual $i$, the number of social ties (i.e., neighbors) in the union network who report using U-Bridge in the past 12 months. We also construct equivalent count measures for the four network types that make up the union network (i.e., friends, family, lenders, and problem solvers). When examining *fractional* threshold models, these variables are measured as the share of adopting neighbors among $i$'s social ties.

While network ties account for *social* influence, we also account for *spatial* influence by using geographic information system data we collected on respondents' household location. The variable *geography* is a spatial lag that counts the number of adopters within the village besides node $i$, and it assigns less weight to those who reside farther away from that node.[8]

We also collect individual-level control variables that likely affect the usage of U-Bridge. These include respondents' sex, age, education, wealth, leadership position, pro-sociality, political participation, and attendance in U-Bridge's inception meeting. We describe how those covariates are measured in SI Section 2.3. At the village level, we compute network measures associated with the social diffusion process, such as density, mean path length, and clustering. We also construct several standard predictors of political participation derived from the 2014 census. Descriptive statistics by village type are shown above in Table 1.

## Estimating Peer Effects

We estimate peer effects, conditional on village type (high/low uptake), using a spatial autoregressive (SAR) model, where the probability of adoption depends on some function of the adoption choice of one's neighbors. Consider individual $i$ embedded in village network $g$ with type $h_g = 1$ if village $g$ is high uptake, and 0 otherwise. $N_i(g)$ is the set of $i$'s neighbors on $g$, and $y_i$ is $i$'s outcome, equal to 1 if $i$ adopts, and 0 otherwise; $y_{N_i(g)}$ is the vector of outcomes of her neighbors, $x_i$ a vector of control variables, and $\epsilon_{ig}$ an error term. Formally:

$$y_{ig} = \beta_{0g} + f(y_{N_i(g)})\beta_1 + h_g f(y_{N_i(g)})\beta_2 + x_i^T\beta_3 + \epsilon_{ig}. \tag{6}$$

We examine both absolute and fractional threshold models with and without controls. In the first case, $f(y_{N_i(g)}) = \sum_{j \in N_i(g)} y_j$ is the number of adopting neighbors. In the second case, $f(y_{N_i(g)}) = \frac{1}{|N_i(g)|}\sum_{j \in N_i(g)} y_j$ is the share of adopting neighbors. For ease of interpretation, we consider linear probability models estimated using ordinary least squares (unless otherwise noted). Conservatively, we account for village-level heterogeneity by using village fixed effects ($\beta_{0g}$). The coefficient $\beta_1$ captures peer effects in low-adoption villages, and $\beta_1 + \beta_2$ is the effect of peers in high-adoption villages.[9] According to our model, $\beta_1 = 0$ and $\beta_1 + \beta_2 > 0$. Due to the small number of clusters, we use bootstrapped standard errors clustered at the village level with 10,000 replicates. In all estimation figures, we report both 95 and 90% confidence intervals using thin and thick bars, respectively.

Whether using the *number* of adopting neighbors (Table 2, columns 1–2) or the *share* of adopting neighbors (Table 2, columns 3–4), adoption of the U-Bridge platform increases with the adoption decisions of one's social ties, *but only in high-uptake villages.*

According to the baseline absolute threshold model (column 2), the likelihood of using U-Bridge increases by 2.9 percentage points for every adopting neighbor in high-uptake villages, which is a 45% increase relative to the mean adoption rate in those villages. Conversely, an additional adopting neighbor increases the likelihood of using U-Bridge by 0.4 percentage points in low-uptake villages, which is substantively small and statistically insignificant. In the baseline "fractional" threshold (column 4), moving from no adopting neighbor to 100% adopting neighbors increases the likelihood of adoption by 3.1 percentage points in low-uptake villages and 28 percentage points in high-uptake villages. These effects, of course, must be calibrated against the data: 32% of respondents have no ties to an adopter, and among those connected to at least one adopting neighbor, the mean *share* of adopting peers is 15%. Moving from no adopting neighbor to 15% adopting neighbors increases the likelihood of adoption by 0.5 percentage points in low-uptake villages, and by 4 percentage points in high-uptake villages.

## Robustness Checks

To check the robustness of our peer effects variability finding, we relax assumptions made in the above analysis and otherwise alter the modeling strategy. To test that the average difference in peer effects between high- and low-uptake villages is not driven by a small number of outliers, we supplant Equation (6) which pools low- and high villages by using a Bayesian multilevel model with random intercepts and slopes (see SI Section 5.1.1 for additional details). With $n_{ig} = \sum_{j \in N_i(g)} y_j$ as the number of adopting neighbors that $i$ has in village $g$, the SAR model in Equation (6) becomes

$$y_{ig} = \beta_{0g} + \beta_{1g} n_{ig} + x_i^T\beta_2 + \epsilon_{ig}, \tag{7}$$

where $\beta_{0g}$ and $\beta_{1g}$ are, respectively, random intercepts and slopes. Figure 5 shows the estimated random slopes ($\beta_{1g}$) in each village. Confirming the pooled specification, almost all high-uptake villages show large, significant peer effects. Conversely, in all low-uptake villages, peer effects are small in magnitude and not significantly different from zero.

[8]With $y_i \in \{0, 1\}$ $i$'s outcome and $d_{ij}$ the distance between $i$ and $j$, the spatial influence (*geography*) is $\text{geo}_i = \sum_{j \neq i} \frac{y_j}{\log d_{ij}}$.

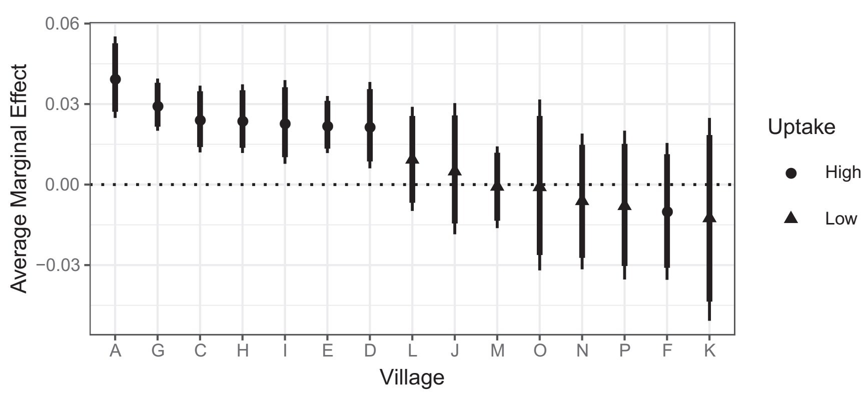
[9]A main effect for high-uptake villages is dropped since it is subsumed by the village-level fixed effects.

**TABLE 2 Adoption of U-Bridge**

| | Dependent Variable: Adopt | | | |
|---|---|---|---|---|
| | Parsimonious (1) | Baseline (2) | Parsimonious (3) | Baseline (4) |
| # Adopting Neighbors ($\beta_1$) | 0.017** | 0.004 | | |
| | (0.007) | (0.006) | | |
| # Adopting Neighbors × High Uptake ($\beta_2$) | 0.021*** | 0.025*** | | |
| | (0.006) | (0.006) | | |
| % Adopting Neighbors ($\beta_1$) | | | 0.102* | 0.031 |
| | | | (0.060) | (0.059) |
| % Adopting Neighbors × High Uptake ($\beta_2$) | | | 0.322*** | 0.244*** |
| | | | (0.107) | (0.087) |
| Degree | 0.002*** | 0.001** | 0.004*** | 0.003*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $\beta_1 + \beta_2$ | 0.038*** | 0.029*** | 0.424*** | 0.276*** |
| Controls | - | √ | - | √ |
| Observations | 3,019 | 3,019 | 3,019 | 3,019 |
| $R^2$ | 0.141 | 0.278 | 0.117 | 0.263 |

*Note*: Absolute threshold Models (Models 1–2) have weakly better fit than fractional threshold Models (Models 3–4). Model 2 is our preferred specification. See the subsection "Estimating Peer Effects" for details about estimation. *p < .1; **p < .05; ***p < .01.

**FIGURE 5 Average Marginal Effect of One Adopting Neighbor on Adoption by Village**




*Note*: Estimates from a multilevel model suggest that save for village F, high-uptake villages have large, significant peer effects. Low-uptake villages have small, statistically insignificant peer effects. Village B is omitted because its sample size is too small.

We further test the robustness of our findings to alternative modeling strategies. First, we use stronger definitions of adoption by increasing the threshold used to define an adopter from having sent at least one message in the past 12 months to thresholds of three and five messages (SI Table 12). Second, we fit logistic regressions instead of linear probability models (SI Table 13). Third, we test whether our results are sensitive to dropping village B, which has a smaller number of respondents as compared to other villages (SI Table 14). Fourth, we explore whether our main results are sensitive to using directed instead of undirected ties (SI Table 15). Fifth, we test sensitivity to the type of ties used to construct the network (SI Table 16). In all cases, we find a strong, positive relationship between the number (or share) of adopting neighbors and one's adoption choice in high- but not low-adoption villages. These checks and their results, which strengthen our confidence in the robustness of our core peer effects finding, are described in greater detail in SI Section 5.

Identifying peer effects causally in observational settings is notoriously difficult. We identify two important sources of confounding and perform three additional analyses that show that our results are likely to be causal (SI Section 5.2). One possible confounder is that the initial encouragements to adopt a technology might be endogenous: even in the absence of social learning, two connected individuals may exhibit similar behavior as a result of homophily or because they are subject to related unobserved shocks. We address this challenge by generalizing An's (2016) instrumental variable approach to multiple peers (SI Section 5.2.1). We leverage an instrument that pushes alter $j$ to adopt, but only affects ego $i$'s adoption decision through $j$'s influence. Our instrument is the distance from one's household to the location of the meeting introducing U-Bridge, as individuals located closer to the venue are more likely to attend the meeting, learn about the program, and in turn adopt the technology.

A second possible confounder is that exposure to peer influence is endogenous to one's network position. Individuals with more central network positions are more likely to be exposed to peer influence since they have more neighbors, or neighbors who are themselves more central. We address this by comparing individuals who share similar network positions (SI Section 5.2.2). Although our main specification controls for one's degree, we push such comparisons further by controlling for degree more flexibly, and for a variety of other centrality scores. Finally, we address both issues jointly using matching (SI Section 5.2.3). Following Aral, Muchnik, and Sundararajan (2009), we construct a matched sample in which villagers share similar individual and network characteristics but differ in the number of their peers who adopted the technology. This procedure alleviates both concerns, since individuals in the matched sample have similar likelihoods of being exposed to treatment owing to their observable individual and network characteristics.

## Discounting and Enforcement Hypotheses

The fact that peer effects are only present in high-uptake villages does not tell us about villages' capacity to enforce truthful communication. Even though (in equilibrium) we cannot observe such capacity directly, we explore several testable implications of this part of our argument. First, villages should differ in the extent to which peer effects foster adoption above and beyond what can be explained by differences in the extent to which peer effects foster diffusion of information about the platform's existence. Indeed, our model emphasizes that such differential effects owe to agents' processing differently the information they obtain from their peers about the technology, and not to differences in their likelihood of obtaining such information in the first place.

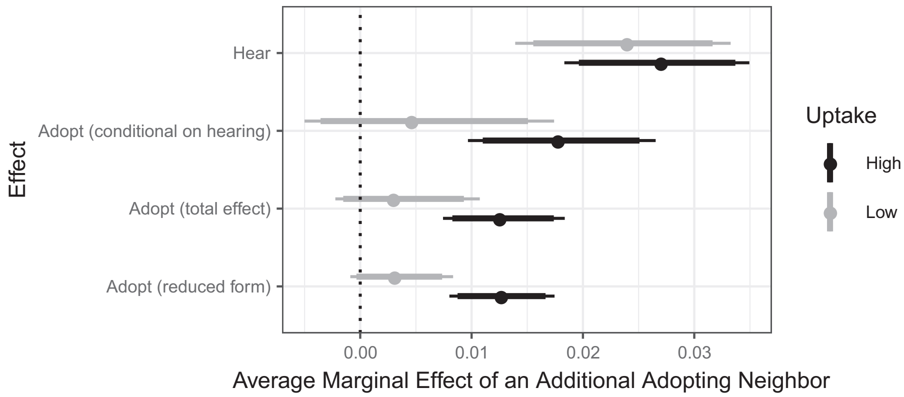
Building on Larson, Lewis, and Rodrıguez (2017), we estimate a two-stage selection model in which we model separately the social process of hearing about an innovation and that of adopting it conditional on hearing. Figure 6 reports those estimates (results in tabular form are reported in SI Section 6.1). In both high- and low-uptake villages, peers affect the likelihood of hearing about the technology. Yet, only in high-uptake villages do peers also affect the likelihood of *adoption conditional on hearing* about the new PCT. As a result, peers only affect the likelihood of adoption in high-uptake villages.

Second, our *discounting* hypothesis states that in the absence of truthful communication, agents discount peers' signals. If it is the case that high-uptake villages enforced truthful communication whereas low-uptake villages did not, then villagers should discount positive signals from peers in low-uptake villages, but they should not in high-uptake villages. We test this by estimating separately the effect of peers who state being satisfied by the platform and those who do not (models reported in SI Table 24, columns 1 and 2). In high-uptake villages, a satisfied peer increases the likelihood of adoption by 2.6 percentage points (p-value = 0.010). In low-uptake villages, a satisfied peer increases the likelihood of adoption by 0.1 percentage points (p-value > .10).

Third, according to our enforcement hypothesis, truthful communication emerges when the cost of misrepresentation is high. Since *strong ties* are more likely to be associated with higher costs of lying, they should be more conducive to peer effects. We thus disaggregate all network relations into simple ties ($i$ shares a single type of relationship with $j$) and complex ties ($i$'s relationship with $j$ is based on more than one of four types of ties). We reestimate our absolute threshold model, first comparing the effect of a complex tie to that of any simple tie, then to that of each kind of simple tie. Consistent with our expectation, we find that peer effects are stronger for complex ties than for simple ties (SI Table 25, column 1). Notably, among simple ties, friendship and family ties are more influential than ties with lenders and problem solvers (column 2).

Fourth, we should observe truthful communication, and hence peer effects, when formal or informal

**FIGURE 6 Selection Model with Hearing**




*Note*: Average marginal effect of an adopting neighbor on hearing (first stage) and of adopting conditional on hearing (second stage). Peers impact hearing about the innovation in both high- and low-uptake villages. Yet, only in high-uptake villages do peers also affect the likelihood of adopting it conditional on hearing about it. The selection model's total effect matches the estimate from a reduced-form logistic regression (see SI Section 6.1 for details about estimation).

institutions are strong enough to impose high costs of lying. While we cannot say with certainty which specific institutions these are, we test several alternatives derived from past work. One possible institution is concentrated leadership, which improves communities' ability to coordinate around shared goals and to sanction (potential) defectors. Coordination and social sanctioning, in turn, may be instrumental in helping communities enforce truthful communication in the face of positive externalities. Other theoretically driven (potential) mediators we test include ethnic and religious homogeneity and (mean) pro-sociality.

To explore the mediating role of concentrated leadership, we conducted a modified public goods game in all 16 villages. Following conventional practice, villagers were given an opportunity to contribute to the village any share of their survey participation remuneration, and the research team matched those contributions. In our version of the public goods game, villagers were asked to name which individual they would like to handle funds on behalf of the village, regardless of whether that individual holds a formal leadership position. We measure leadership concentration as a Herfindahl index based on these responses. We rerun our multilevel specification allowing the coefficient on the number of adopting neighbors to be a function of not only the village-level random component $b_{1g}$, but also $z_g$, which is the village-level leadership concentration.

$$y_{ig} = \beta_{0g} + (\beta_{1g} + z_g^T \gamma) n_{ig} + x_i^T \beta_2 + \epsilon_{ig}. \quad (8)$$

We find that leadership concentration is likely a mediator of the relationship between peer effects and adoption. The coefficient on the interaction is 0.077, 95% CI [0.028, 0.126], suggesting that the more concentrated leadership is, the stronger peer effects are (SI Table 26).[10] This finding is consistent with the idea that leadership concentration supports truthful communication in the face of externalities.

We do not find support for other alternative mediators (SI Table 27). First, we examine ethnic and religious homogeneity, measured by Herfindahl indexes calculated from the 2014 census. Ethnic homogeneity does not mediate the effect of peers, but peer effects are somewhat larger in villages that are more religiously homogeneous. Next, we examine pro-sociality, measured as village-level mean contributions to dictator and public goods games. Here again, the interaction effect is significant. Peer effects are significantly larger in villages with higher levels of pro-sociality. However, unlike leadership concentration, high-uptake villages do not exhibit greater religious homogeneity or pro-sociality compared to low-uptake villages (SI Figure 7). As such, these mediators do not help explain cross-village variation in the strength of peer effects. Since we have only 16 villages, these results, while consistent with our theoretical framework, should be viewed primarily as an invitation for further research.

[10] Our findings are robust to different definitions of leadership concentration.

**TABLE 3 Descriptive Statistics about Posteriors**

| Variable | Sample | High uptake | Low uptake | Δ |
|---|---|---|---|---|
| Government responsiveness | $3.214^{***}$ | $3.22^{***}$ | $3.205^{**}$ | 0.015 |
| Δ Government responsiveness | $3.394^{***}$ | $3.378^{***}$ | $3.416^{***}$ | −0.038 |
| Government capacity | $3.939^{***}$ | $3.955^{***}$ | $3.916^{***}$ | 0.039 |
| Δ Government capacity | $3.652^{***}$ | $3.618^{***}$ | $3.699^{***}$ | −0.081 |
| Quality of education | $3.091^{*}$ | 3.041 | $3.158^{**}$ | −0.118 |
| Quality of health clinics | $2.694^{**}$ | $2.589^{**}$ | $2.835^{**}$ | −0.246 |
| Quality of access to water | $2.175^{***}$ | $2.32^{**}$ | $1.982^{***}$ | 0.338 |
| Quality of roads | $2.327^{***}$ | $2.246^{***}$ | $2.435^{***}$ | −0.19 |

*Note*: Each variable is measured on a 1–5 scale. Rows that start with Δ ask for perceived variation in the past 12 months, with 3 corresponding to no change. The columns "Sample," "High Uptake," and "Low Uptake" test for whether the mean value is different from 3. The column Δ tests for whether the difference between high- and low-uptake villages is significantly different from zero. For each test, standard errors are clustered at the village level. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Evidence of signal discounting in low-uptake villages, combined with the demonstrated variation in peer effects, give us confidence the observed divergence in uptake cannot be fully explained by collective action problems arising in some villages and not others. Earlier, we argued that this scenario was unlikely because citizens had similar grievances and faced an equally responsive government. These additional results—namely, variation in peer effects and discounting of signals—cannot be accounted for by collective action problems alone. They are, however, consistent with our theory.

## Examining Other Model Implications to Determine the Scenarios That Explain Divergences in Outcomes

We now turn to two additional model implications and attempt to determine which of the various scenarios outlined in Figure 3 best explain the divergence in outcomes observed in these villages. We first establish that the state of the world was likely the same in high- and low-uptake villages.

Recall that the state of the world captures whether the government is both responsive to citizens' demands and capable of addressing them (high state), or is not responsive or incapable of addressing citizens' requests (low). We provided evidence above that the Arua local government was equally responsive in high- and low-uptake villages.

Two additional results bolster the claim that the state of the world was likely the same in high- and low-uptake villages. First, examining the quality of education services using administrative data and unannounced audits conducted at baseline and endline in public schools in the study area (SI Table 7), we find little improvement in either high- or low-uptake villages. Additionally, when present, improvements are not significantly different between high- and low-uptake villages. Second, we elicit survey respondents' (posterior) beliefs about their local government's capacity, will to respond to citizen complaints, and evaluation of the quality of public services. We find that high- and low-uptake villages have indistinguishable posteriors (Table 3); their evaluations of government responsiveness (row 1), government capacity (row 3), and the quality of public services (rows 5–8) show no significant differences. At the time we conducted the survey, citizens had 2 years of experience with U-Bridge, sufficient time for information about the program to overwhelm prior beliefs. If both high- and low-uptake villages reach the same conclusions about the state of the world, then the state of the world must be the same in those villages.

We further establish that high-uptake villages likely had higher priors than low-uptake villages. Table 3 provides suggestive evidence: Low-uptake villages seem to have updated their beliefs on government capacity and responsiveness to a higher extent than high-uptake villages (rows 2 and 4), although the difference is insignificant. Since high- and low-uptake villages converged to the same posterior beliefs, it follows that low-uptake villages had lower priors to begin with. Moreover, according to our model, if priors are low, then initial adoption will be low, whereas it will be high if priors are high. This is apparent when examining patterns of adoption over time (Figure 2), where high-uptake villages send significantly more messages (per 100 residents) than low-uptake villages in the first few months after launch. Differential priors are also suggested by patterns of meeting attendance. Although meeting attendees had similar characteristics in high- and low-uptake villages, there were fewer such attendees in low-uptake villages, which may reflect lower interest in the platform, possibly driven by lower priors (SI Table 9).

We can now determine which scenario delineated by our theory (Figure 3) best describes adoption patterns in high- and low-uptake villages. We have established that (1) high-uptake villages likely enforced truthful communication, whereas low-uptake villages likely did not; (2) high-uptake villages likely had more optimistic priors than low-uptake villages; and (3) the state of the world was likely the same in high- and low-uptake villages. Patterns of adoption over time are, in turn, increasing and then decreasing for high-uptake villages. They increase slightly in low-uptake villages and then decrease, but are globally low (Figure 2). These patterns suggest that citizens initially put greater weight on government responsiveness when trying to determine the state of the world, and they inferred that the state was high because district authorities were very reactive to incoming messages (response rate was about 90%). We believe that during the first year of the program, high-uptake villages thought they were in quadrant H1, whereas low-uptake villages were in quadrant H4 of Figure 3. However, a year into the program, citizens likely put more weight on government capacity, ultimately inferring the state of the world is low based on the lack of improvement in the quality of public services. As such, a year into the program, high-uptake villages likely moved from quadrant H1 to L1, whereas low-uptake villages moved from quadrant H4 to L4.

## Conclusion

In this study, we explain variation in the adoption of new PCTs. Since new technologies are costly and their benefits are uncertain, potential users rely on the experience of early adopters in their social network. We argue that the diffusion process of any new technology is governed by the extent to which the benefits of adoption depend on other agents' actions, and we develop a model that clarifies how and why the information-sharing process within a network could differ for goods with substantial positive externalities compared to those with minimal externalities. A key contribution of this study is therefore to offer a new, more general theory of technology adoption that, unlike previous work, can better explain why many new technologies for *political* engagement fail to take off.

Adopting a new technology for political communication belongs to a broader class of political actions, like joining a protest, that are costly and characterized by externalities and uncertainty about the returns to taking action, as well as potential for learning and communication about those returns in a social network. Past work on such forms of participation has focused on how networks facilitate *coordination* (Steinert-Threlkeld 2017). In this study, we highlight that the role of networks in facilitating political action is (also) crucially mediated by the quality, or truthfulness, of communication.

To understand whether and when peer effects will facilitate adoption of a new technology with positive externalities, we must assess the extent to which communities have mechanisms for enforcing truthful communication about the costs and benefits of the technology. The same logic extends to describing the life span of a social movement and participation in protests. Our model sheds new light on an old insight: the first movers—also referred to as the core (Steinert-Threlkeld 2017)—will be the most committed individuals in a group. However, it also suggests that first movers will tend to distort the private information they acquired at the protest and exaggerate the probability of success of future protests in order to foster future participation. Therefore, members of a movement will discount information from first movers, leading to suboptimal actions in the short and perhaps even the long run, unless they are part of a group with strong internal norms governing truth telling (Eubank and Kronick 2019).

In our case, examining adoption patterns of a new technology for political communication in rural Uganda, we show that peer effects, and hence technology diffusion, emerge in some but not all villages. Some villages were unable to establish truthful communication, and the reports of early adopters were discounted by their peers. Our sample of villages is not sufficiently large to establish with confidence exactly when and how villages overcome the impediments to truthful information sharing, but we show suggestive evidence that concentrated leadership and strong social ties might facilitate diffusion. By contrast, we find little evidence that the structure of the network itself is consequential. These findings offer promising avenues for further research. In addition, we leave for future work a thorough treatment of the possibility of negative externalities.

We also contribute to an expanding literature exploring the effects of social networks on political behavior. Existing work focuses mostly on well-established forms of engagement like voting (Siegel 2013). We investigate the role of social networks in the adoption of novel forms of political engagement, where there is higher uncertainty over costs and benefits of participation, and thus peer effects and communication are arguably more important. Finally, by situating our study in a low-income country, we join others (e.g., Cruz, Labonne, and Querubin 2017; Larson and Lewis 2017) in moving beyond the prevailing focus on networks and political behavior in a small number of industrial democracies.

New PCTs cannot improve governance if they go unused. Social networks play an important role in social diffusion processes, but political technologies are unique. Externalities and information-sharing barriers explain PCTs' low rates of adoption, and variation in adoption rates across communities.

## References

Acemoglu, Daron, Asuman Ozdaglar, and Ercan Yildiz. 2011. "Diffusion of Innovations in Social Networks." In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference.* IEEE 2329–34. Orlando, FL, USA.

An, Weihua. 2016. "On the Directionality Test of Peer Effects in Social Networks." *Sociological Methods & Research* 45(4): 635–50.

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): 21544–49.

Bénabou, Roland, and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." *Quarterly Journal of Economics* 126(2): 805–55.

Bhatti, Zubair K., Jody Zall Kusek, and Tony Verheijen. 2014. *Logged On: Smart Government Solutions from South Asia.* Washington, DC: World Bank Publications.

Blair, Graeme, Rebecca Littman, and Elizabeth Levy Paluck. 2019. "Motivating the Adoption of New Community-Minded Behaviors: An Empirical Test in Nigeria." *Science Advances* 5(3): eaau5175.

Centola, Damon. 2015. "The Social Origins of Networks and Diffusion." *American Journal of Sociology* 120(5): 1295–1338.

Centola, Damon, and Michael Macy. 2007. "Complex Contagions and the Weakness of Long Ties." *American Journal of Sociology* 113(3): 702–34.

Cruz, Cesi, Julien Labonne, and Pablo Querubin. 2017. "Politician Family Networks and Electoral Outcomes: Evidence from the Philippines." *American Economic Review* 107(10): 3006–37.

Engelmann, Dirk, and Martin Strobel. 2004. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments." *American Economic Review* 94(4): 857–69.

Eubank, Nicholas, and Dorothy Kronick. 2019. "Friends Don't Let Friends Free Ride." Working paper. https://bit.ly/2JizjF8.

Fearon, James D., and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90(4): 715–35.

Kolaczyk, Eric D. 2009. *Statistical Analysis of Network Data.* New York, NY: Springer.

Larson, Jennifer M., and Janet I. Lewis. 2017. "Ethnic Networks." *American Journal of Political Science* 61(2): 350–64.

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. http://bit.ly/2DwT2Lt.

Peixoto, Tiago, and Micah L. Sifry, eds. 2017. *Civic Tech in the Global South: Assessing Technology for the Public Good.* Washington, DC: World Bank and Personal Democracy Press.

Rueda, Miguel R. 2015. "Buying Votes with Imperfect Local Knowledge and a Secret Ballot." *Journal of Theoretical Politics* 27(3): 428–56.

Siegel, David A. 2013. "Social Networks and the Mass Media." *American Political Science Review* 107(4): 786–805.

Sjoberg, Fredrik M., Jonathan Mellon, and Tiago Peixoto. 2017. "The Effect of Bureaucratic Responsiveness on Citizen Participation." *Public Administration Review* 77(3): 340–51.

Steinert-Threlkeld, Zachary C. 2017. "Spontaneous Collective Action: Peripheral Mobilization during the Arab Spring." *American Political Science Review* 111(2): 379–403.

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Section 1**: Glossary of network concepts
**Section 2**: Additional information on the setting
**Section 3**: Additional information on the puzzle
**Section 4**: Model
**Section 5**: Robustness checks and causal inference
**Section 6**: Additional evidence on the empirical implications of the model

# 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.

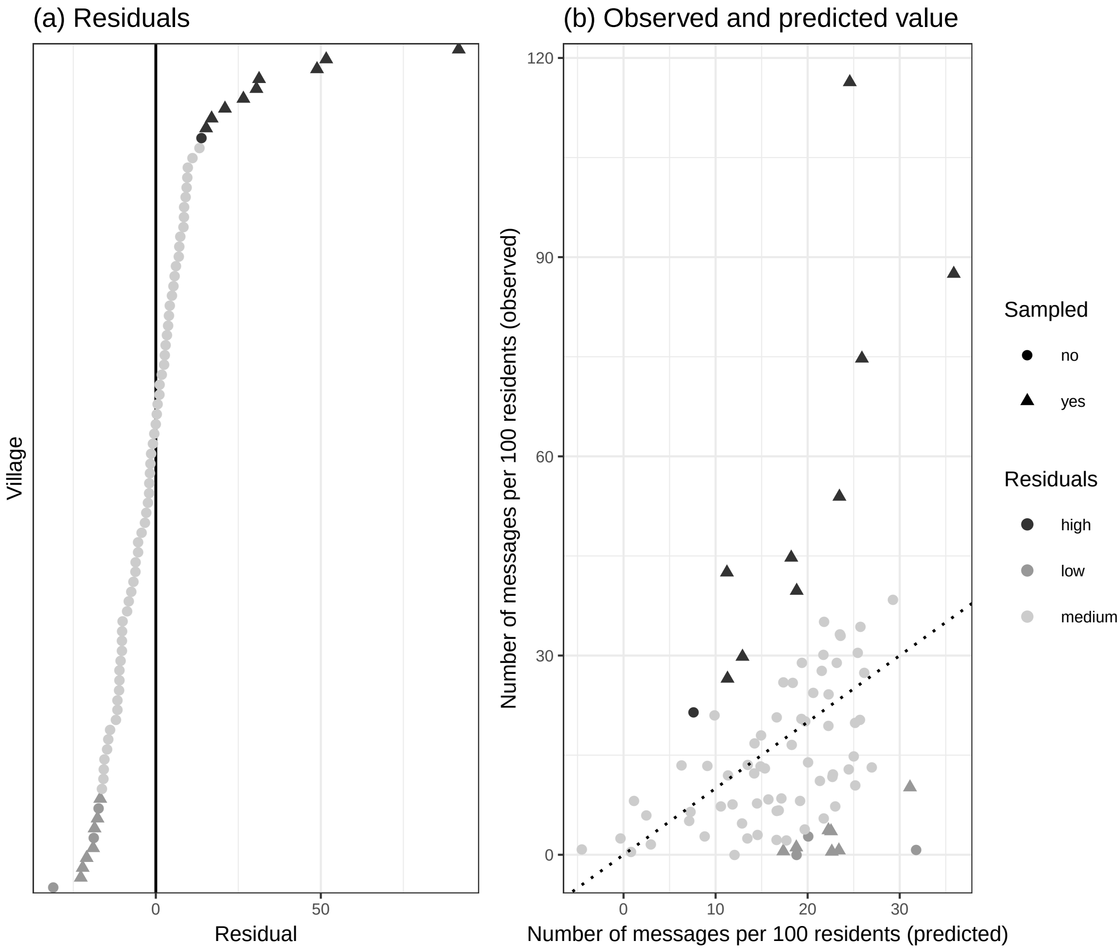



Figure 1: **Residuals of village selection regression.** Colors report the 10 villages with highest and lowest residuals. Triangles report, out of these 20 villages, the ones that were included in the sample.

## 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.

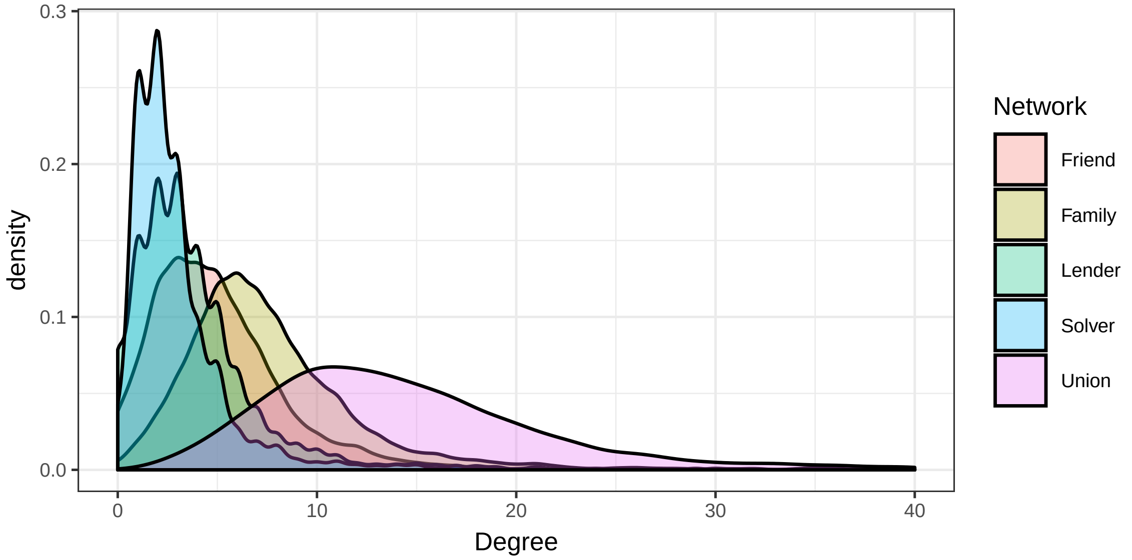



Figure 2: Degree distribution by network type.

## 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 | $\Delta$ |
|---|---|---|---|---|---|---|---|---|
| **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 $\Delta$ 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) | $\Delta$ |
| 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.

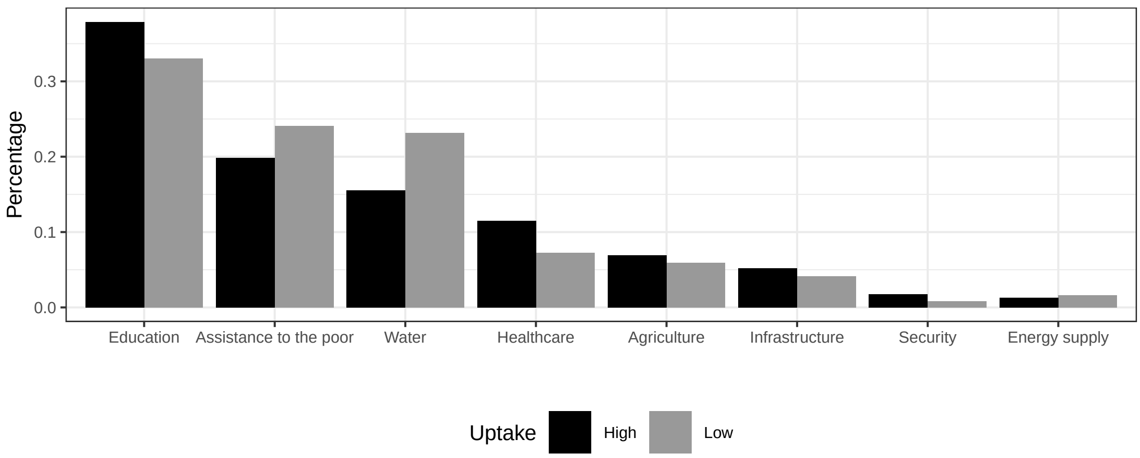



Figure 3: **Policy priorities across high- and low-uptake villages for all surveyed individuals.** High- and low-uptake villages have comparable policy priorities: their orderings of priorities are similar. The finding is inconsistent with an alternative explanation according to which high-uptake villages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

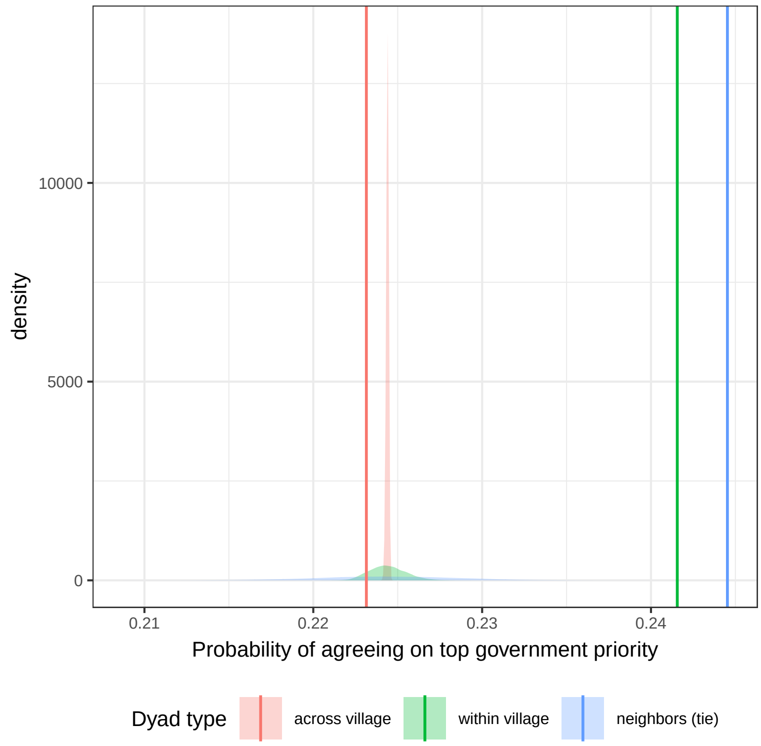



Figure 4: **Agreement on government priorities.** Probability that two individuals agree if they live in different villages, in the same village, or are connected on the union network. Bars represent observed values. Shaded areas represent the distribution of these statistics in a permutation test where 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).

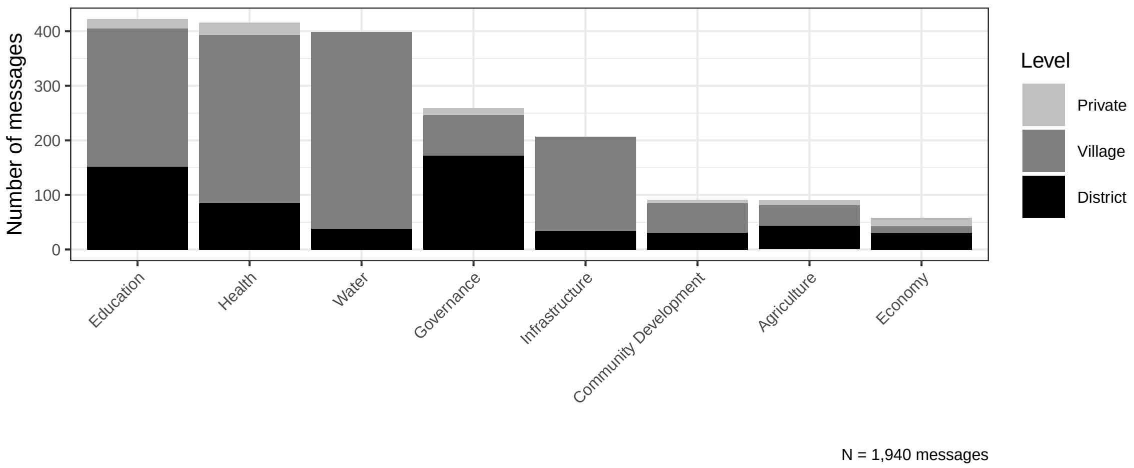



Figure 5: **Messages by topic and level.** Irrespective of the topic, the overwhelming majority of messages point out problems at the district or village level, instead of private problems.

### 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 $\sigma_0$ 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 $\varphi:\mathcal{I}_{i2}^{g'}(y_{i0},y_{-i0})\to\mathcal{I}_{i2}^{g}(y_{i0},y_{-i0})$ 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 $\sigma$. 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 $\sigma$. 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

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

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

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

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 $V^\sigma_{i1}(\mathcal{I}_{i1}(y_{i0})) \equiv \mathbb{E}_\sigma[v(M^{s*}_i)^\gamma+u(y^*_{i2},\theta)^{\gamma^2}|y_{i0}]$ is the Blackwell information value of the subgame that starts at $t=1$ following action $y_{i0}$ under profile $\sigma$. Substituting and solving for $\pi_i$ in the above inequality gives

$$y^*_{i0}=1 \iff \pi_i \geq a^\sigma_{i0} \equiv \frac{[c-p_L]-[V^\sigma_{i1}(\mathcal{I}_{i1}(1))-V^\sigma_{i1}(\mathcal{I}_{i1}(0))]}{p_H-p_L}$$

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 $\sigma$ 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 $\theta$ 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 $\pi_i$ 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 $\sigma$, with some probability, we have $m_{ji} \neq s_j$. Conversely, under $\sigma_0$, we always have $m_{ji} = s_j$. As such, one may *garble* the conditional distribution of $m_{ji}$ under $\sigma_0$ to recover the conditional distribution of $m_{ji}$ under $\sigma$. 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 $\sigma$ 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 $\sigma$ 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 $\sigma_0$ 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 $\sigma_0$, agent $i$ plays lottery $l_0 \in \mathcal{L}_{\mathcal{M}_i}$ such that $\Pr(m = m_0) = 1$. The profile $\sigma_0$ is an equilibrium if and only if

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

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 $\sigma_0$ 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 $\Sigma$ 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

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

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 ($\beta_1$) | 0.002 | |
| | (−0.009, 0.014) | |
| # adopting neighbors × high-uptake ($\beta_2$) | 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 ($\beta_1$) | 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 ($\beta_2$) | 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 ($\beta_1$) | 0.705*** | 0.195 | | |
| | (0.190) | (0.188) | | |
| # adopting neighbors × high-uptake ($\beta_2$) | −0.259 | 0.138 | | |
| | (0.164) | (0.179) | | |
| % adopting neighbors ($\beta_1$) | | | 1.923 | −1.184 |
| | | | (2.839) | (4.247) |
| % adopting neighbors × high-uptake ($\beta_2$) | | | 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 ($\beta_1$) | 0.017** | 0.005 | | |
| | (0.007) | (0.006) | | |
| # adopting neighbors × high-uptake ($\beta_2$) | 0.022*** | 0.025*** | | |
| | (0.006) | (0.006) | | |
| % adopting neighbors ($\beta_1$) | | | 0.102* | 0.034 |
| | | | (0.060) | (0.058) |
| % adopting neighbors × high-uptake ($\beta_2$) | | | 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 ($\beta_1$) | 0.039*** | 0.025*** | | |
| | (0.009) | (0.007) | | |
| # adopting in-neighbors × high-uptake ($\beta_2$) | 0.010 | 0.014* | | |
| | (0.009) | (0.008) | | |
| % adopting in-neighbors ($\beta_1$) | | | 0.107** | 0.085 |
| | | | (0.053) | (0.086) |
| % adopting in-neighbors × high-uptake ($\beta_2$) | | | 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 ($\beta_1$) | 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 ($\beta_2$) | 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 $\epsilon$ an error term. Formally, the problem is that the autoregressive term $My$ is quite possibly correlated with the error term $\epsilon$. To address the issue of endogeneity, An (2016) recommends using an instrument $z$ that is correlated with $y$, but not with $\epsilon$. Using the two stages least squares (2SLS) procedure, we estimate the following models with OLS:

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

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 ($\beta_1$) | 0.016 | 0.008 | 0.004 | | | |
| | (0.014) | (0.011) | (0.008) | | | |
| # adopting neighbors × high-uptake ($\beta_2$) | 0.029** | 0.040*** | 0.025*** | | | |
| | (0.013) | (0.010) | (0.008) | | | |
| % adopting neighbors ($\beta_1$) | | | | 0.147 | −0.062 | 0.031 |
| | | | | (0.151) | (0.117) | (0.069) |
| % adopting neighbors × high-uptake ($\beta_2$) | | | | −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 ($\beta_1$) | 0.008 | 0.003 |
| | (0.005) | (0.006) |
| # adopting neighbors × high-uptake ($\beta_2$) | 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 ($\beta_1$) | 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 ($\beta_2$) | 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.

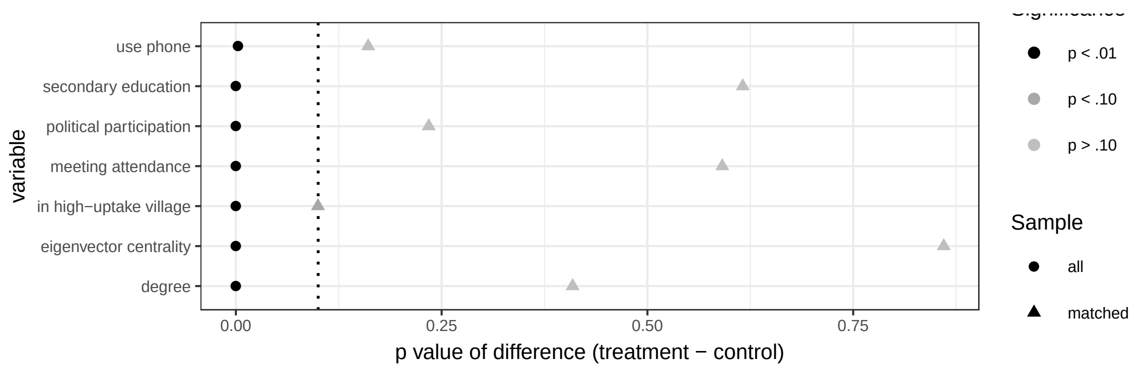



Figure 6: **Covariate balance on dimensions used for matching ($t = 1$).** We report the p-value of the difference in means in the treatment and control group in the full and matched sample, using full matching. Matching achieves balance in degree, phone use, and meeting attendance.

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 $\geq t$ ($\beta_1$) | 0.026** | −0.021 | 0.011 |
| | (0.010) | (0.017) | (0.021) |
| # adopting neighbors $\geq t\times$ high-uptake ($\beta_2$) | 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 ($\beta_1$) | $0.155^{***}$ | | |
| | (0.033) | | |
| # hearing neighbors × high-uptake ($\beta_2$) | 0.0003 | | |
| | (0.031) | | |
| # adopting neighbors ($\beta_1$) | | 0.139 | 0.195 |
| | | (0.152) | (0.188) |
| # adopting neighbors × high-uptake ($\beta_2$) | | 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 ($\beta_1$) | 0.009 | 0.030$^{***}$ | 0.006 |
| | (0.009) | (0.005) | (0.007) |
| # non-satisfied adopting neighbors × high-uptake ($\beta_2$) | | | 0.026$^{***}$ |
| | | | (0.009) |
| # satisfied adopting neighbors ($\beta_3$) | 0.001 | 0.025$^{**}$ | −0.001 |
| | (0.010) | (0.010) | (0.011) |
| # satisfied adopting neighbors × high-uptake ($\beta_4$) | | | 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, $\beta_s$ | 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, $\beta_c$ | 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

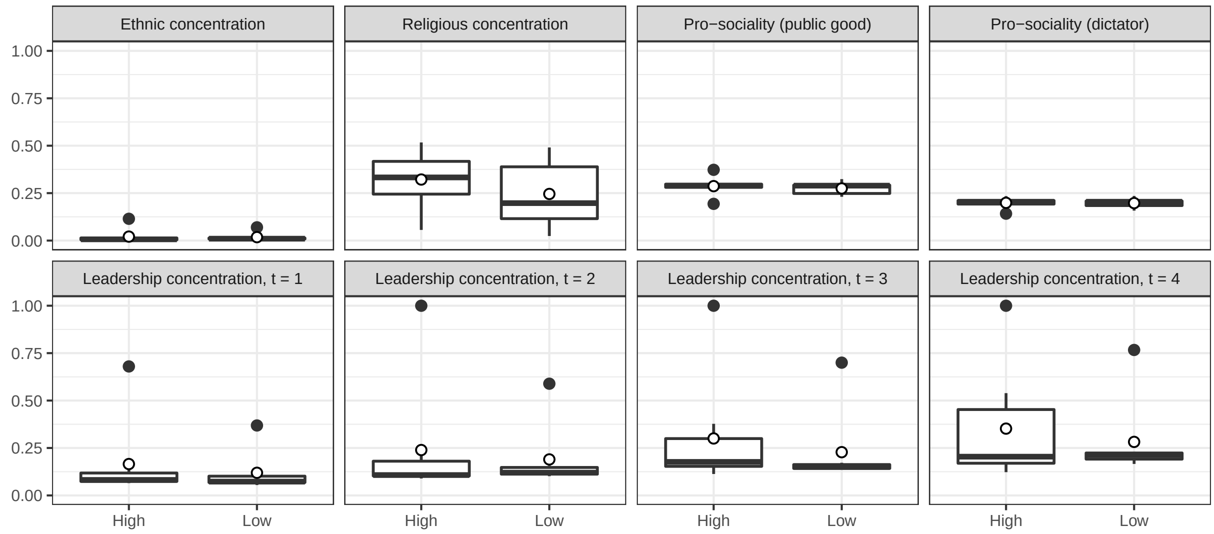



Figure 7: **Mediators in high and low uptake villages.** White points represent the mean. The distribution of mediators is comparable in high and low uptake villages for ethnic and religious concentration and for pro-sociality. High-uptake villages have more concentrated leadership than low-uptake villages.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