# CAN FRIENDS SEED MORE BUZZ AND ADOPTION?

By

Vineet Kumar and K. Sudhir

May 2019

Revised October 2021

COWLES FOUNDATION DISCUSSION PAPER NO. 2178R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

# Can Friends Seed More Buzz and Adoption?

Vineet Kumar
Yale School of Management, 165 Whitney Avenue, New Haven, CT 06511, vineet.kumar@yale.edu

K. Sudhir
Yale School of Management, 165 Whitney Avenue, New Haven, CT 06511, k.sudhir@yale.edu

A critical element of word of mouth (WOM) or buzz marketing is to identify seeds, often central actors with high degree in the social network. Seed identification typically requires data on the full network structure, which is often unavailable. We therefore examine the impact of WOM seeding strategies motivated by the friendship paradox to obtain more central nodes *without knowing network structure* on adoption. Higher-degree nodes may be less effective as seeds if these nodes communicate less with neighbors or are less persuasive when they communicate; therefore whether friendship paradox motivated seeding strategies increase or reduce WOM and adoption remains an empirical question. We develop and estimate a model of WOM and adoption using data on microfinance adoption across 43 villages in India for which we have data on social networks. Counterfactuals show that the proposed seeding strategies are about 15-24% more effective in increasing adoption relative to random seeding . These strategies are also about 5-13% more effective than the firm's leader seeding strategy, and are relatively more effective when we have fewer seeds.

*Key words*: word of mouth, networks, seeding, friendship paradox, product adoption, diffusion

## 1. Introduction

Firm-initiated and consumer-driven word of mouth (WOM) marketing (often referred to as buzz marketing), has received a lot of attention, and has proven effective in increasing adoption across a wide range of products and services. WOM has been examined both theoretically and empirically using a wide range of modeling approaches to understand both the motivations to engage in it and its various impacts (Godes and Mayzlin 2009, Iyengar et al. 2011, Campbell et al. 2017, Berger and Iyengar 2013, Cai et al. 2015).

An important question in WOM marketing is how to choose appropriate seeds. There are a few broad approaches considered in the literature. The first uses network data on connections to identify central individuals (e.g. degree or eigenvector centrality) to provide the most WOM (Tucker 2008, Goldenberg et al. 2009, Libai et al. 2013). Recently, researchers have also tried to combine multiple networks among the same individuals to identify seeds with specific relationship types that might lead to higher adoption (Chen et al. 2017). The second uses characteristics of individuals to identify how opinion leaders can be used to seed networks (Iyengar et al. 2011). Opinion leaders are often highly context-specific and may not span multiple categories, e.g. an opinion leader in fashion might not be so when it comes to consumer electronics or healthcare (King and Summers 1970).

Another approach is to identify seeds based on local network properties and community characteristics to achieve higher diffusion (Yoganarasimhan 2012). There might be tradeoffs in terms of the local structure, where network structures that enable high diversity of content might not be efficient at accelerating the flow of information (Aral and Van Alstyne 2011). The outcome of diffusion might typically be context-dependent, and thus an approach that is broadly applicable and theoretically founded would be helpful.

Broadly, the emphasis in the recent literature has been to improve seed identification using richer and and more comprehensive network data. However, even with easier access to online social networks, data on the right or *relevant network for a particular purpose* is often unavailable. For example, even if one had access to the Facebook social networks of everyone including physicians, the *relevant* physician-to-physician network data for seeding a new drug may be unavailable. Even within a specific context, there are a number of challenges to gathering accurate network data, including the time and effort required to obtain this data (Stark 2018). Moreover, the dynamically evolving nature of connections and relationships requires frequent updating of such data. Social media data, which are relatively easier to access also face the challenge that activity there maybe more of a substitute than a complement to offline or other social interactions and may not be effective in high-involvement applications (Borgatti et al. 2009).

This paper investigates a complementary approach that obviates the need to use detailed network data by introducing WOM seeding strategies *when the relevant network structure information is unavailable*. The strategies leverage the Friendship Paradox to identify *more connected* individuals for seeding irrespective of the underlying network structure.

**Friendship Paradox and Network Seeding Strategies**

Put simply, the friendship paradox can be stated as "On average, your friends have more friends than you do." The paradox is based on a mathematical result that holds independent of network structure, because popular people are always over-represented when averaging over friends (Feld 1991), and its use has been suggested for immunization of networks and sensors on networks (Cohen et al. 2003, Christakis and Fowler 2010). The basic intuition is simple: Suppose we choose an initial node at random (so each node has an equal probability of being selected) and then choose one friend of that node at random. The chosen friend of the initial mode is likely to be more highly connected than average, because by construction a highly connected node will be in the friend set of more people, and therefore more likely to be nominated as a friend. Consider two extreme examples for intuition. First, in a simple hub-spoke network with a central node and several peripheral nodes, which are all connected *only* to the central node. Each node has equal probability of being initially selected, so we are very likely to get a peripheral node. When asked to nominate a friend, all the peripheral nodes can only suggest the central node, who is their only friend. Second, if we have an isolated node without any connections, such a node would never be chosen by anyone on the network as a friend.

The friendship paradox thus suggests potential strategies for sampling higher degree individuals (ones with more friends) in any network, without knowing network structure (Kumar et al. 2018). This sampling approach of choosing a random friend is termed as the "Local Friend" strategy (local friend of friend) and is informationally light in that it only requires access to a set of randomly sampled individuals, and the ability to obtain a random friend from them. The other advantage is that the list of relevant friends from which to sample can be easily adjusted as appropriate for the particular seeding problem at hand. The friendship paradox proof guarantees that individuals with higher than average degree are obtained *in expectation*, allowing for potentially better seeding. Even though one can sample higher degree individuals using these strategies, their use as seeds cannot guarantee greater WOM or product adoption, because the extent to which higher degree individuals communicate with friends in their network about the product is an empirical question. In a recent study, Kim et al. (2015) found that selecting the highest degree nodes did not result in higher adoption relative to random seeding. They also found mixed evidence for the effectiveness of friend nominations across two different categories; the mixed effects could be due to differences in network structure across the villages in their treatment and control groups.[1]

The above discussion motivates our research questions below.

(1) Can friendship paradox based seeding strategies improve WOM and adoption relative to random seeding? Can it improve upon an opinion leader based strategy chosen by the firm? By how much?

(2) Can hybrid approaches leveraging the friendship paradox along with leadership characteristics lead to higher adoption?

(3) How does the extent of initial seeding (proportion of the network seeded) impact *absolute and relative performance* of the strategies?

To address these questions, we empirically model the WOM and product adoption process over networks by allowing for a flexible relationship between degree and WOM. Further, in contrast

[1] It is challenging to control for network structure experimentally since the number of possible networks structures grows exponentially in the number of nodes. For $N = 100$ nodes, there are $2^{\frac{N(N-1)}{2}} \approx 10^{1490}$ possible undirected network structures. There have been recent efforts at evaluating the effectiveness of random and multi-hop stochastic seeding strategies using nonparametric estimation approaches (Chin et al. 2018), where the typical assumption is that seed sets are mapped to outcomes in a fixed manner.

to typical diffusion models, which *assume* that all WOM arises from adopters, our model incorporates WOM from both adopters and non-adopters, which enables us to quantify their relative contribution to WOM.

Estimating such a WOM diffusion model is challenging because the necessary multi-network data is typically unavailable. Most diffusion models are estimated based on one product's time series of adoption through one market (or social network). Further, the original seeding is typically unobserved, and even if observed it is often not possible to identify the effect of different seeding without multiple diffusion paths across similar networks. Finally, the impact of WOM might be misidentified in the presence of advertising (Van den Bulte and Lilien 2001), attributing to WOM what was actually achieved by advertising.

In this paper, we are able to address each of these challenges through data on one product (microfinance) adoption across 43 independent and relatively isolated village social networks in India. The firm's seeding across the different villages leads to exogenous variation in network position and characteristics of seeds, which aids in identifying the impact of seeding. Also, there was no advertising or promotion activity by the firm that would confound WOM effects, which is known to bias estimates of impact (Van den Bulte and Lilien 2001).

Based on the estimates, we simulate counterfactuals on WOM and product adoption across these villages as a function of alternative WOM seeding strategies. Finally, we compare the effectiveness of the friendship paradox based Local Friend and hybrid seeding strategies relative to Random and Opinion Leader based strategies. We find that the friendship paradox based strategies obtain a significant (24%) improvement over the Random strategy, and even the Leader strategy used by the firm. The Local Friend strategy used in conjunction with the Leader strategy resulting in a hybrid provides a marginal improvement over either. Finally, strategies that are informationally more demanding like Top Degree and Top Diffusion perform better than other strategies, obtaining close to 44% improvement over the random baseline.

## 2. Data

We use panel data on the diffusion and adoption of microfinance across households belonging to 43 villages in Southern India in combination with rich network data on the social connections among the households within each village. The data was collected and described in Banerjee et al. (2013). The microfinance firm identified opinion leaders based on leader and social criteria in each village prior to entry and seeded information about the microfinance product among these individuals first. Tables 1 and 2 provide the summary statistics of village social networks, household characteristics and microfinance adoption in the villages. In Table 2, the statistics are for giant component within each network.[2]

From Table 1, we see that households have an average of more than 4 individuals. Taken as averages across villages, 92% of households have electricity, but only 29% of households have latrines. We note the relatively lower variation in the number of people relative to rooms or beds across the households.

**Table 1 Summary Statistics of Village Networks**

| Statistic | Mean | SD | Min | Max |
|---|---|---|---|---|
| *Degree:* | | | | |
| Mean | 9.66 | 1.64 | 6.82 | 13.59 |
| Standard Deviation | 7.09 | 1.32 | 5.18 | 11.02 |
| Minimum | 1.00 | 0.00 | 1.00 | 1.00 |
| Maximum | 39.72 | 13.01 | 23.00 | 90.00 |
| Mean of Leaders | 12.94 | 2.59 | 8.88 | 18.82 |
| *Household Characteristics:* | | | | |
| Number of Households | 223.21 | 56.17 | 114.00 | 356.00 |
| People in Household | 4.48 | 0.54 | 3.34 | 5.83 |
| Rooms in Household | 2.31 | 0.41 | 0.75 | 2.94 |
| Beds in Household | 0.88 | 0.46 | 0.29 | 2.27 |
| Proportion of Households with Electricity | 0.92 | 0.11 | 0.23 | 0.98 |
| Proportion of Households with Latrines | 0.29 | 0.16 | 0.02 | 0.91 |
| Proportion of Households with Leaders (%) | 12.55 | 3.16 | 7.00 | 21.00 |

*Note: Unit of analysis is a village network (N=75)*

**Table 2 Summary Statistics of Adoption (%)**

| Statistic | Mean | SD | Min | Max |
|---|---|---|---|---|
| All Households | 19.38 | 8.16 | 8.00 | 45.00 |
| Leader | 24.71 | 12.64 | 3.57 | 53.85 |
| Followers | 18.68 | 8.19 | 7.30 | 43.71 |
| Electrified Households | 19.01 | 8.38 | 7.34 | 45.12 |
| Non-electrified Households | 20.72 | 13.26 | 0.00 | 50.00 |
| Latrine Households | 14.69 | 9.17 | 0.00 | 36.36 |
| Non-latrine Households | 21.85 | 9.89 | 7.03 | 51.25 |

*Note: Unit of analysis is a village network (N=75)*

Table 1 summarizes the characteristics of the village social networks. There is considerable variation in the extent of relationships among households. Each village contains on average 212

[2] The giant component of a network is the largest connected component of the network, excluding isolated nodes.

households. Across villages, the mean degree (connections) of households is around 9, The mean of the standard deviation of degree for households at the village level is large at around 7.1, with the minimum and maximum also reflecting wide variation. The mean degree of opinion leaders is higher than the average and close to the maximum of average degree across villages. We observe that opinion leaders have a much higher degree than average. Finally, note that the variation in degree across villages and among households within villages provides the identifying variation to estimate the diffusion model proposed in Section 3 below.

The primary performance comparison in our study is the adoption of microfinance by households across the villages. The adoption in the data are detailed in Table 2. We find that 19.2% of households adopt microfinance, with significant variation across the villages. Opinion leaders are more likely to adopt than followers, perhaps a feature of the information propagation chosen by the firm, which targeted these leaders in each village. Adoption is correlated with Household characteristics; electrified households are less likely to adopt compared to non-electrified, and households with a latrine are less likely to adopt than those without. Broadly, these results suggest that microfinance is needed by households at the bottom of the pyramid in emerging markets.

## 3. Model and Estimation

We use a model of WOM and product adoption across a social network. Using network terminology, households are *nodes* and connections between them are *edges*. Households need to be informed about the product in order to adopt. Households who are informed communicate with their neighbors probabilistically, even if they have not adopted. We build upon the model of Banerjee et al. (2013), with key adaptations required to study our research question related to the friendship paradox. First, we allow the WOM probability from a node to differ by degree, reflecting the idea that WOM effort may depend on this factor. Second, we allow the WOM probability from a node to differ for those identified as "leaders" by the firm. Banerjee et al. (2013) allow the probability of WOM to depend on adoption status but not on the number of connections (degree) or for "leaders". Our extensions are specifically motivated by the strategies examined here. If we choose their specification, we will obtain a probability of communication that depends on adoption status but not on the number of connections (degree). Recall that our approach based on Friendship Paradox obtains higher degree nodes than average. Thus, we chose a conservative approach, allowing for the idea that whereas high degree nodes may be better due to their degree, they *might* also be less likely to communicate with their friends. If we did not account for that, then we could be biasing the results *in favor of the friendship paradox strategy*. Similarly, accounting for differences in WOM among firm designated "leaders" is critical to assess the effectiveness of leader strategies.

**Baseline Model**

*Word of Mouth Communication:* WOM occurs in the network when a household receives information (only) from its *informed* neighbors. We allow WOM probability $p^s(D)$ to depend on adoption status $s$ and degree $D$.

$$p^s(D) = q^s_{min} + (q^s_{max} - q^s_{min}) \left[ \frac{D - D_{min}}{D_{max} - D_{min}} \right] \tag{1}$$

Thus, $q^s_{min}$ represents the WOM probability for a node with minimum degree ($D = D_{min}$), whereas $q^s_{max}$ represents the WOM probability for the highest degree node ($D = D_{max}$) and adoption status is denoted $s$. These quantities are based on the minimum and maximum degrees across all networks. Both parameters depend on the adoption status $s \in \{NA, A\}$ of the node, with $NA$ indicating "Not Adopted" and $A$ indicating "Adopted." The specification in Banerjee et al. (2013) is a special case of this model when $q_{min} = q_{max} = q$, such that WOM is independent of degree. Nodes continue communicating with neighbors in periods after they become informed.

*Adoption:* When a household becomes aware of the product at time $t$, the household's decision of whether to adopt, $y \in \{0, 1\}$, is modeled as a standard logit choice with observed heterogeneity. The utility of household $i$ from adoption and non-adoption is:

$$\begin{aligned} u_i(y=1) &= \beta_0 + \beta X_i + \epsilon_{i,1} \\ u_i(y=0) &= \epsilon_{i,0} \end{aligned} \tag{2}$$

$X_i$ represents the vector of leader characteristics of household $i$, $\beta$ the vector of coefficients, and $\epsilon_{i,s}$ are distributed as Type I EV random variables.

After a node becomes *informed* either as an initial seed or through a neighbor, further WOM from others does not impact the likelihood of adoption. Thus, WOM is purely informational rather than persuasive in this baseline.While the baseline model provides a useful benchmark, it leads to the question of whether there are more complex decision processes for communication and adoption, which we examine next.

**Endorsement or Persuasion**

In the endorsement or persuasion model, (termed "complex contagion" by Centola and Macy (2007)), likelihood of adoption varies based on whether WOM is received from more friends. Following Banerjee et al. (2013), the utility of adoption is:

$$u_i(y=1) = \beta_0 + \beta X_i + \lambda F_{it} + \epsilon_{i,1} \tag{3}$$

where $F_{it}$ is the fraction of neighbors who have informed $i$ about microfinance and $\lambda$ is the endorsement parameter. The utility of non-adoption remains unchanged.

**Leader Effects**

Leaders selected as seeds by the firm may have unobserved individual characteristics (leadership) that lead to higher probability of WOM relative to non-leaders, over and above their higher degree. Further, firms may have provided specific information to their selected leader seeds, which may make their WOM more effective.[3] To capture such differences, we extend the baseline model to allow for differential probability of WOM for leaders:

$$p_i^s(D) = q_{min}^s + (q_{max}^s - q_{min}^s)\left[\frac{D - D_{min}}{D_{max} - D_{min}}\right] + q_\ell \mathbf{1}[i \in Leaders] \tag{4}$$

Thus, if leaders are especially effective in spreading WOM, we would find the parameter $q_\ell$ to be positive, whereas a negative value would indicate leaders are less effective than others.[4]

**Nonlinear Effect of Degree**

Finally, we allow WOM likelihood to be nonlinear in degree by allowing a quadratic effect, which can also capture potential non-monotonicity with respect to degree.

$$p^s(D) = q_{min}^s + (q_{max}^s - q_{min}^s)\left[\frac{D - D_{min}}{D_{max} - D_{min}}\right] + q_Q\left[\frac{D - D_{min}}{D_{max} - D_{min}}\right]^2 \tag{5}$$

where $q_Q$ represents the parameter corresponding to the quadratic term.

We examine a number of different models, summarized in Table 3. Overall, we have 8 specifications. The first 4 models have no endorsement or persuasion effect (denoted by superscript $\mathbf{E=0}$). In $\mathbf{M_1^{E=0}}$, the WOM probability does not depend on degree. This model is identical to the model in Banerjee et al. (2013). In $\mathbf{M_2^{E=0}}$, the WOM probability depends on degree as detailed in Section 3. $\mathbf{M_3^{E=0}}$ incorporates a differential effect for leaders to the prior model specification. $\mathbf{M_4^{E=0}}$ allows for nonlinear relationship between WOM probability and degree with a quadratic function. The next four models are identical to the first four, but with an endorsement effect (denoted by superscript $\mathbf{E=1}$).

**Table 3 Summary of WOM Model Components**

| | $\mathbf{M_1^{E=0}}$ | $\mathbf{M_2^{E=0}}$ | $\mathbf{M_3^{E=0}}$ | $\mathbf{M_4^{E=0}}$ | $\mathbf{M_1^{E=1}}$ | $\mathbf{M_2^{E=1}}$ | $\mathbf{M_3^{E=1}}$ | $\mathbf{M_4^{E=1}}$ |
|---|---|---|---|---|---|---|---|---|
| Endorsement | × | × | × | × | ✓ | ✓ | ✓ | ✓ |
| Degree-dependent WOM | × | ✓ | ✓ | ✓ | × | ✓ | ✓ | ✓ |
| Leader Differential WOM | × | × | ✓ | ✓ | × | × | ✓ | ✓ |
| Nonlinear Effect: WOM and Degree | × | × | × | ✓ | × | × | × | ✓ |

[3] Our model does not distinguish between incidence of WOM and its effectiveness, but so long as both those effects do not change in the counterfactual, the strategy comparisons remain valid.

[4] We note that since all initial seeds are "leaders," it aids the leader fixed effect identification as any impact of leader fixed effect will be stronger in the initial periods and can be therefore identified off the adoption trajectory. More details about the identification of the Leader Fixed Effect is detailed in §EC.2.

## Estimation

The model estimation proceeds in three steps similar to Banerjee et al. (2013), with specific differences. Our estimation is detailed in Supplement §EC.3. Note that we use optimization algorithms for estimation rather than grid search. Here we provide a high level description of the three steps.

*Step 1: Adoption Process* – We estimate the adoption process parameters $\beta$ with a logistic regression using the adoption decisions of only the initially seeded individuals based on equation (2).

*Step 2: WOM Process* – We estimate the WOM process parameters $(q^{NA}_{min}, q^{NA}_{max}, q^{A}_{min}, q^{A}_{max})$ as well as endorsement ($\lambda$), leader effect ($q_\ell$) and quadratic effect ($q_Q$) using the method of simulated moments (MSM). We use the same set of cross sectional moments used in Banerjee et al. (2013), supplemented by time series moments, all listed in Table 4. Overall, the moments capture key aspects of diffusion within a network, both globally over the entire network and locally across connections. The first moment is global, matching overall adoption levels in the network. Moments 2-4 are local moments that fit household level adoption as a function of adoption characteristics of their neighbors, and help identify communication probabilities for non-adopters and adopters respectively. Moments 5 and 6 are also local moments in that they capture covariance in adoption between a household and its first and second degree neighbors respectively.[5] Next, we include time series moments that have not been used in Banerjee et al. (2013). Moments 7-9 characterize the temporal trajectory of adoption within villages, which helps us in the identification of the leader fixed effect, as detailed in §EC.2. We detail how each moment informs the estimation of each parameter, i.e. the sensitivity of parameter estimates to each of the cross sectional and time series moments based on the approach of Andrews et al. (2017) in §EC.5.7.

The objective function for the parameter vector $\theta$ is defined as in Banerjee et al. (2013):

$$S(\theta) = \left(\frac{1}{S}\sum_{s=1}^{S}\left[m^S(\theta) - m^D\right]'\right)\mathbf{W}\left(\frac{1}{S}\sum_{s=1}^{S}\left[m^S(\theta) - m^D\right]\right) \tag{6}$$

where $m^S(\theta)$ represents the vector of model simulated moments, $m^D$ denotes the vector of data moments. $W$ is the weighing matrix, which can either be estimated with a two–stage approach or be set to be the identity matrix to obtain consistent estimates. The estimator is then defined as:

$$\hat{\theta} = \arg\min_{\theta} S(\theta) \tag{7}$$

*Step 3: Standard Errors* – We estimate the standard errors using a block-bootstrap resampling procedure of sampling with replacement, treating each network as a block.

[5] We provide precise specification of the moments and the rationale for using them in §EC.3.

**Table 4 List of Moments**

| | |
|---|---|
| 1. | Proportion of seeds adopting |
| 2. | Proportion of households with no adopting neighbors who have adopted |
| 3. | Proportion of neighbors of adopting seeds who have adopted |
| 4. | Proportion of neighbors of non-adopting seeds who have adopted |
| 5. | Covariance between a household's adoption and average adoption of its first degree neighbors |
| 6. | Covariance between a household's adoption and average adoption of its second degree neighbors |
| 7,8,9 | Cumulative adoption upto time $t = 1, 2, 3$ (Time series moments) |

## 4. Results

We first detail the results from the adoption model in Table 5. We find that the number of beds in the household and the rooms per person are negatively associated with adoption probability, whereas access to latrine in the home and beds per person has a positive impact. The estimates are not only consistent with the idea that microfinance is typically used by relatively poor households without access to traditional banking services, but that the poorest households are not the biggest adopters.[6]

**Table 5 Adoption: DV: Microfinance Adoption (1=yes, 0=no).**

| Variable | Estimate | SE |
|---|---|---|
| Constant | $-1.210^{***}$ | (0.322) |
| Rooms | 0.007 | (0.085) |
| Beds | $-0.283^{**}$ | (0.143) |
| Electricity | 0.156 | (0.123) |
| Latrine | $0.179^{**}$ | (0.080) |
| Rooms per person | $-1.023^{***}$ | (0.392) |
| Beds per person | $1.147^{*}$ | (0.656) |
| Log Likelihood | $-603.093$ | |
| *Note:* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | |

Table 6 reports the estimates for the 8 WOM models. We use the SMM (or GMM) J-statistic (a measure of fit) for model selection, which is a standard measure of model fit for GMM models and is asymptotically distributed as $\chi^2$ under the null (Hansen 1982). Based on the Sargan's J-test using the J-statistic, we cannot reject the null that the model is valid. Given that the J-statistic is lowest for Models $\mathbf{M_2^{E=0}}$ (without endorsement) and $\mathbf{M_2^{E=1}}$ (with endorsement), we use these as our primary specifications to interpret parameters and for counterfactual analysis. [7]

[6] We have examined a variety of adoption models in Section EC.5.5 and the results presented here represent the model with best fit (lowest AIC).

[7] We also consider a large number of other models to test various other specifications of the data generating process. We mention a few here. First, we allowed for a "broadcast process" to consider the effect of an initial village meeting by seeds to communicate to all households. Second, we considered the potential for a "leader certification" effect to allow for an incremental effect of leader endorsement on adoption. Third, we allowed for opinion leader seeds to be chosen by an occupation, auch that leader effects are conferred on members of the occupation. None of these model

**Table 6 Model Estimates**

| Parameter | Symbol | Model Specification: Estimates (Standard Errors) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *No Endorsement* | | | | *With Endorsement* | | | |
| | | $\mathbf{M_1^{E=0}}$ | $\mathbf{M_2^{E=0}}$ | $\mathbf{M_3^{E=0}}$ | $\mathbf{M_4^{E=0}}$ | $\mathbf{M_1^{E=1}}$ | $\mathbf{M_2^{E=1}}$ | $\mathbf{M_3^{E=1}}$ | $\mathbf{M_4^{E=1}}$ |
| Non-adopter lowest degree | $q_{min}^{NA}$ | 0.041 | 0.036 | 0.061 | 0.061 | 0.041 | 0.041 | 0.061 | 0.056 |
| | | (0.001) | (0.004) | (0.011) | (0.033) | (0.003) | (0.004) | (0.019) | (0.023) |
| Non-adopter highest degree | $q_{max}^{NA}$ | 0.041 | 0.036 | 0.051 | 0.051 | 0.041 | 0.038 | 0.051 | 0.038 |
| | | (0.001) | (0.012) | (0.0001) | (0.016) | (0.003) | (0.007) | (0.068) | (0.062) |
| Adopter lowest degree | $q_{min}^{A}$ | 0.341 | 0.400 | 0.386 | 0.366 | 0.362 | 0.339 | 0.396 | 0.406 |
| | | (0.012) | (0.086) | (0.028) | (0.003) | (0.009) | (0.014) | (0.047) | (0.059) |
| Adopter highest degree | $q_{max}^{A}$ | 0.341 | 0.348 | 0.338 | 0.339 | 0.362 | 0.356 | 0.309 | 0.326 |
| | | (0.012) | (0.106) | (0.034) | (0.073) | (0.009) | (0.029) | (0.052) | (0.048) |
| Leader Effect | $q_\ell$ | – | – | -0.022 | -0.021 | – | – | -0.021 | -0.014 |
| | | | | (0.012) | (0.046) | | | (0.007) | (0.01) |
| Quadratic Effect | $q_Q$ | – | – | – | 0.029 | – | – | – | 0.016 |
| | | | | | (0.053) | | | | (0.03) |
| Endorsement | $\lambda$ | – | – | – | – | -0.036 | -0.007 | -0.021 | -0.034 |
| | | – | – | – | – | (0.046) | (0.076) | (0.067) | (0.07) |
| *J-Statistic* ($\times 10^{-6}$) | | 2.905 | 2.837 | 3.268 | 3.179 | 3.088 | 2.967 | 3.333 | 3.337 |

In Models $\mathbf{M_1^{E=0}}$ and $\mathbf{M_1^{E=1}}$, grayed out parameters are not estimated since $q_{min}^s = q_{max}^s$.

We first interpret the parameter estimates of the preferred model specifications $\mathbf{M_2^{E=0}}$ and $\mathbf{M_2^{E=1}}$. We begin with the case of no endorsement. First, the WOM probability for adopters is much higher than that of non-adopters, by an order of magnitude ($q_{min}^A \gg q_{min}^{NA}$). Next, we examine degree dependence. For non-adopters, the WOM probability does not depend on household degree ($q_{min}^{NA} \approx q_{max}^{NA}$), so that low-degree households are as likely as high-degree households to communicate with each of their network neighbors. For adopters, however, the high-degree households are less likely to communicate with each of their peers relative to low-degree households ($q_{max}^A < q_{min}^A$). Yet, high-degree households communicate more overall since they have more connections.

From $\mathbf{M_3^{E=0}}$ and $\mathbf{M_4^{E=0}}$, we find no differential effect of leaders; the parameter $q_\ell$ is small, negative and not statistically significant, implying that leaders neither communicate more nor are more effective. For the quadratic effect, we do not find $q_Q$ to be statistically significant.

Finally, we find no evidence of an endorsement or persuasion effect, estimated through parameter $\lambda$ in models $\mathbf{M_1^{E=1}}$ and $\mathbf{M_4^{E=1}}$. Across all four models, the persuasion effect is small in magnitude, negative in sign and not statistically significant. This is consistent with Banerjee et al. (2013) (their specification corresponds to $\mathbf{M_1^{E=1}}$). For the other parameters, the estimates are similar to the models without the endorsement effect. In the supplement §EC.5, we consider a number of other

specifications fit the data better than $\mathbf{M_2^{E=0}}$. The counterfactual performance under *all* of the models are provided in §EC.4.3). We also discuss additional model fit metrics in §EC.3.1, evaluating both in-sample and out-of-sample fit for different model specifications.

additional models of the data generating process.[8] In all cases, our baseline model indeed does fit best relative to other models, so we use that as the primary specification for the counterfactuals.

## 5. Counterfactuals

We use counterfactuals to evaluate various seeding strategies based on Friend, Leader, Hybrid and Network Information categories described in Table 7. Within the Friend category, we examine the impact of the Local Friend strategy, which samples on friends of randomly chosen network nodes (households) to obtain seeds. In Leader, we examine both the "(Firm's) Leader" strategy, using the original leaders that were designated for seeding by the microfinance firm, and "Like Leader," which chooses as seeds leader-like nodes who have similar network positions as leaders.

We use three dimensions to measure network position: degree, eigenvector and power centrality (Bonacich 1987). This helps evaluate whether the impact of seeding is due to the network position or due to the differential impact by individual characteristics of leaders. Hybrid strategies combine the features of sampling on friends along with information on opinion leaders. We examine two different hybrid strategies: choosing a random *Friend of Leader* household (weak hybrid) or choosing a random *Leader Friend of Leader* household (strong hybrid).

We also evaluate two network information benchmarks, Top Degree and Top Diffusion. Unlike the above strategies, these network strategies require strong knowledge on who is connected to whom (network structure) or the degree distribution, i.e. the number of connections of each node. If highly connected nodes are likely to be better in accelerating adoption, the idea is to choose from the set of highly connected nodes in the Top Degree strategy. Top Diffusion is an approach proposed by Banerjee et al. (2013) to identify nodes with high centrality for the purpose of information diffusion. It requires the social network structure (adjacency matrix), but does not require knowledge of the parameters of the diffusion process. Seeds are randomly chosen from the set of top 15% of nodes for both top degree and top diffusion strategies. Further details about the strategies, including informational requirements, are provided in Section EC.4.2.

We use the estimated parameters from $\mathbf{M_2^{E=0}}$ for the counterfactual simulations below. In the Supplement, we provide a comparison of the counterfactual results of all the different model specifications summarized in Table 3. We set seeding level at 1% of the number of households in the village; therefore number of households seeded varies across villages as a function of village populations. We examine the sensitivity of the results to different seeding levels (0.5%, 1%, 5%) in Section 5.1 below.

[8] We consider a benchmark where there are only broadcasts (§EC.5.1), but no communication through networks. In §EC.5.2, we consider a model where there is an initial broadcast by seeds, to model an initial village meeting described in Banerjee et al. (2013). We also consider the case where leaders may be present outside the initial seed set (§EC.5.3), and where leaders may have specific certification ability (§EC.5.4). Specifically, we show in Figure EC.3 that there is reasonable probability that seeds chosen by any of our counterfactual strategies overlap with the seeds chosen by the firm.

**Table 7 Seeding Strategies and Implementation**

| **Category** | **Strategy** | **Implementation Procedure (for each of $m$ seeds)** |
|---|---|---|
| Friendship | Local Friend | Select node at random from list. Obtain one randomly chosen friend of node as a seed. |
| Leader | (Firm's) Leader | Select node from list of leaders |
| | Like Leader | Select leader node $\ell$ at random. Select the non-leader node most similar to $\ell$ in terms of network properties. |
| Hybrid | Friend of Leader (Weak Hybrid) | Select a random **leader** from list of leaders. Obtain one randomly chosen friend of this leader as a seed. |
| | Leader Friend of Leader (Strong Hybrid) | Select a random **leader** from list of leaders. Obtain one randomly chosen friend who is also a **leader** to be seed. |
| Network Information | Top Degree | Select a seed node at random from the list of top (Top 15%) degree (most connected) nodes . |
| | Top Diffusion | Select a seed node at random from the list of top (Top 15%) diffusion nodes (proposed by Banerjee et al. (2013), and defined in Table EC.5). |

We evaluate seeding effectiveness in terms of proportion of informed households and adoption generated by the seeding strategies as the performance measure. Table 8 reports the aggregate statistics on the proportion of households informed about the microfinance service and the proportion adopting microfinance. The improvement for Local Friend over Random is about 23.7%, while the improvement over Random for Leader is about 14.5%. We also find that the Hybrid strategy Friend of Leader performs the best with a 24.2% improvement over Random, suggesting that the two broad approaches of leveraging network structure (using friendship paradox) and leadership or other demographic characteristics (using Leader indicator) can be combined to achieve higher performance. However, we note that using the Local Friend strategy alone without any information about the network structure or leader information can generate much of this performance benefit. Overall, the Local Friend and Hybrid strategies do better than the Leader strategy without data on network structure, suggesting that they are viable approaches to seeding WOM with unknown networks.

As we might expect, the network information strategies, which require global knowledge of the network structure lead to much greater adoption than the above strategies. Interestingly, the top degree approach performs better than the diffusion centrality based approach, although the difference is minimal. It's likely that for these relatively small village networks, the overlap in seed sets among the top degree and top diffusion strategies is quite high, hence the similarity. In larger networks, e.g. Twitter we might see larger differences between them. Overall, we observe that the Local Friend strategy is able to obtain a little more than 50% of the improvement that the network information strategies obtain above the random strategy baseline. This finding characterizes the tradeoff between the amount of information required and the effectiveness of the strategy in driving adoption. In cases where it is impractical to obtain the *relevant* network information, the Local Friend or hybrid strategies could be profitably used as an alternative.

Table 8: Comparison of Strategies (1% seeding)

| **Strategy** | **Informed (%)** Mean | SD | **Adopted (%)** Mean | SD | **ΔInformed(%)** over Random | **ΔAdopted(%)** over Random |
|---|---|---|---|---|---|---|
| Random | 47.83 | 42.30 | 9.66 | 28.40 | — | — |
| Local Friend | 57.19 | 44.08 | 11.95 | 31.59 | 19.55 | 23.73 |
| (Firm's) Leader | 53.79 | 43.60 | 11.05 | 30.40 | 12.44 | 14.46 |
| Like Leader | 53.10 | 43.32 | 10.90 | 30.14 | 11.00 | 12.90 |
| Hybrid Strategies: | | | | | | |
| Friend of Leader | 57.36 | 43.91 | 11.99 | 31.55 | 19.91 | 24.19 |
| Leader Friend of Leader | 54.53 | 43.40 | 11.27 | 30.58 | 14.00 | 16.74 |
| Network Information Strategies: | | | | | | |
| Top Degree | 64.41 | 43.97 | 13.98 | 33.99 | 34.66 | 44.77 |
| Top Diffusion | 63.79 | 44.07 | 13.82 | 33.79 | 33.35 | 43.15 |

Note: Parameter Estimates from model $\mathbf{M}_{\mathbf{2}}^{\mathbf{E=1}}$ used for counterfactuals.

We report the pairwise comparison between strategies Table 9. The Local Friend strategy is better than Random and leads to improved adoption in most of the villages. The Local Friend strategy also outperforms the Leader strategy across most villages (88.37%). The (Firm's) Leader strategy does worse than Random in about 12% of of the villages. The weak hybrid Friend of Leader strategy is also better than random in all villages, but the strong hybrid Leader Friend of Leader actually performs worse than random in about 5% of the villages. This implies that *it matters how the hybrid strategy is implemented*, and whether the condition of leadership is required for not just the initial node but also for the nominated friend. The results suggest reduced effectiveness of seeding when we require that the nominated friend also be a leader. Finally, we note that the Like Leader strategy is the most similar in performance to the Leader strategy, as expected.

### 5.1. How does Extent of Seeding Impact Performance of Strategies?

The purpose of word-of-mouth marketing is to choose a small number of seeds to help spread information about a product or service. We summarize in Table 10 how the performance of the seeding strategies varies with the proportion of nodes seeded, at 0.5%, 1%, and 5% of nodes seeded. For full results across all model specifications, see Supplement Section EC.4.3.

We define the performance metric as leverage, in terms of how well a proposed seeding strategy $s$ performs relative to the Random strategy (whose leverage is 1 by definition):

$$Leverage(s) = \frac{\text{\# Households Adopting under Strategy } s}{\text{\# Households Adopting under Random Strategy}}$$

**Table 9 Pairwise Comparison of Strategies (1% seeding)**

| | Local | Leader | Like Leader | Friend of Leader | Leader Friend of Leader | Top Degree | Top Diffusion |
|---|---|---|---|---|---|---|---|
| Random | 97.67 | 88.37 | 86.05 | 100.00 | 95.35 | 100.00 | 100.00 |
| Local Friend | | 18.60 | 18.60 | 46.51 | 32.56 | 100.00 | 95.35 |
| (Firm's) Leader | | | 44.19 | 86.05 | 53.49 | 100.00 | 100.00 |
| Like Leader | | | | 90.70 | 55.81 | 100.00 | 100.00 |
| Friend of Leader | | | | | 16.28 | 100.00 | 93.02 |
| Leader Friend of Leader | | | | | | 100.00 | 97.67 |
| Top Degree | | | | | | | 41.86 |
| Top Diffusion | | | | | | | |

Note: Number in cell indicates % of villages where **column** strategy achieves higher adoption than **row** strategy.

**Table 10 Leverage for Counterfactual Strategies**

| | | No Endorsement | | | With Endorsement | | |
|---|---|---|---|---|---|---|---|
| Strategy | *Seeding at:* | 0.50% | 1.00% | 5.00% | 0.50% | 1.00% | 5.00% |
| Local Friend | | 1.26 | 1.18 | 1.05 | 1.28 | 1.24 | 1.05 |
| (Firm's) Leader | | 1.14 | 1.08 | 1.03 | 1.13 | 1.15 | 1.03 |
| Like Leader | | 1.12 | 1.10 | 1.03 | 1.13 | 1.13 | 1.03 |
| Hybrid Strategies: | | | | | | | |
| Friend of Leader | | 1.26 | 1.20 | 1.05 | 1.27 | 1.24 | 1.05 |
| Leader Friend of Leader | | 1.17 | 1.13 | 1.02 | 1.18 | 1.17 | 1.02 |
| Network based Strategies: | | | | | | | |
| Top Degree | | 1.53 | 1.40 | 1.07 | 1.55 | 1.45 | 1.09 |
| Top Diffusion | | 1.52 | 1.40 | 1.07 | 1.57 | 1.43 | 1.07 |

The following observations are noteworthy. First, the (Firm's) Leader strategy always outperforms the Random strategy and the Local Friend strategy always outperforms the Leader. Thus, our main results hold across the range of seeding proportions examined. Second, the weak hybrid strategy dominates all the others, whereas the strong hybrid consistently underperforms the Local Friend strategy. Third, Like Leader performs very similar to leader, indicating that performance of the leader strategy is not driven by the differential leader effects, but rather the network position of leaders. Finally, leverage for all strategies decreases as the number of seeds increases, implying that their performance differential is greater under highly constrained circumstances, e.g., when product samples are limited or the seeding requires intensive education or interaction.

## 6. Conclusion

We estimate a model of network-mediated WOM and product adoption and evaluated the effectiveness of alternative seeding strategies that leverage the friendship paradox. The proposed friendship paradox based strategies, which are *informationally light* and require little knowledge of network structure significantly improve WOM seeding and product adoption relative to not just random seeding, but also relative to the firm's opinion leader seeding. Specifically, we find a 15-24% improvement with Local Friend seeding in both information spread and adoption compared with Random, and about 5-13% improvement over the Firm's Leader seeding (based on membership of selected occupations). Further, Local Friend seeding is better than Random across multiple villages with varying network structures, whereas the Leader strategy can be worse than Random in a significant number of village networks. We note that this result is stochastic and based on expected performance. However, it should be clear that seeding strategies that use detailed network information can improve adoption relative to the Local Friend strategy.

We find that the relative advantage of both Local Friend and hybrid strategies relative to the Random strategy is inversely related to the proportion of nodes seeded. Thus, when we have few seeds, these strategies become even more advantageous in expectation. This result is practically useful in cases where the target population is large, and the seeding is costly in monetary terms, or challenging in terms of time constraints or other operational limitations.

The question addressed here deserves further investigation across a wide range of application networks. As discussed in the introduction, Chin et al. (2018) do not find support for the benefits of seeding using the friendship paradox while like our paper, Kim et al. (2015) find support for it. We speculate that one reason that Chin et al. (2018) do not find support could be because the degree range is 1 to 5. In contrast, in our data, node degree ranges from 1 to a maximum of 39; hence the chance of finding higher degree nodes through the friend strategy is greater. This is also consistent with the experimental evidence in Kim et al. (2015) about the role of degree variance in the value of friend based seeding strategies. Another aspect that is worth mentioning is that our model is a parametric specification, whereas the identification logic of leader fixed effect is nonparametric. Although we have estimated a wide variety of models to demonstrate robustness, with any parametric model, we cannot fully rule out all forms of mis-specification.

As we discussed in the introduction, our data across multiple village networks with explicit knowledge of who were chosen as initial seeds has many advantages for studying the current seeding problem. Yet, an important limitation is that we do not observe the information transmission and this is modeled as a latent process driving adoption. Hence the results are potentially impacted by the specific models of information transmission. We address this issue by estimating many alternative models of information transmission and demonstrate that our key counterfactual claims are robust. Similarly, even though the seeding process is observed, the firm always seeded on "leaders" based on a certain set of occupations. Though we show that leader fixed effects can be identified even with the currently available data, it would be useful to examine the value of additional variation provided by explicitly randomizing with leader seeds in some villages and non-leaders in others. It would also be useful in future work to replicate our results in data settings where the incidence of information transmission and even its content is observable, and alternative seeding processes are used in the data. If information transmission can be observed, it will lead to richer insights about how information diffuses through the network under various seeding strategies and ultimately impacts adoption.

There are important issues that we leave to future research. First, it would be useful to explore the potential tradeoff in terms of cost and time in using our informationally light seeding strategies versus investing in identifying (even limited) network information prior to seeding. Second, rather than use opinion leaders, it may be useful to seed individuals nominated by others as "gossipers" to assess their impact on diffusion and higher overall adoption (Stephen and Lehmann 2016, Banerjee et al. 2014). Finally, it would be useful to consider whether seeding approaches proposed here need to be adapted for highly asymmetric networks, where directional ties become significant (Ben Sliman and Kohli 2018).

## References


Andrews I, Gentzkow M, Shapiro JM (2017) Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics* 132(4):1553–1592.

Aral S, Van Alstyne M (2011) The diversity-bandwidth trade-off. *American Journal of Sociology* 117(1):90–171.

Banerjee A, Chandrasekhar AG, Duflo E, Jackson MO (2013) The diffusion of microfinance. *Science* 341(6144):1236498.

Banerjee A, Chandrasekhar AG, Duflo E, Jackson MO (2014) Gossip: Identifying central individuals in a social network. Technical report, National Bureau of Economic Research.

Bass FM (1969) A new product growth for model consumer durables. *Management science* 15(5):215–227.

Ben Sliman M, Kohli R (2018) Asymmetric relations and the friendship paradox. *Columbia Business School Research Paper* (18-73).

Berger J, Iyengar R (2013) Communication channels and word of mouth: How the medium shapes the message. *Journal of Consumer Research* 40(3):567–579.

Bonacich P (1987) Power and centrality: A family of measures. *American Journal of Sociology* 92(5):1170–1182, URL `http://dx.doi.org/10.1086/228631`.

Borgatti SP, Mehra A, Brass DJ, Labianca G (2009) Network analysis in the social sciences. *Science* 323(5916):892–895.

Cai J, De Janvry A, Sadoulet E (2015) Social networks and the decision to insure. *American Economic Journal: Applied Economics* 7(2):81–108.

Campbell A, Mayzlin D, Shin J (2017) Managing buzz. *The RAND Journal of Economics* 48(1):203–229.

Centola D, Macy M (2007) Complex contagions and the weakness of long ties. *American journal of Sociology* 113(3):702–734.

Chen X, van der Lans R, Phan TQ (2017) Uncovering the importance of relationship characteristics in social networks: Implications for seeding strategies. *Journal of Marketing Research* 54(2):187–201.

Chin A, Eckles D, Ugander J (2018) Evaluating stochastic seeding strategies in networks. *arXiv preprint arXiv:1809.09561* .

Christakis NA, Fowler JH (2010) Social network sensors for early detection of contagious outbreaks. *PloS one* 5(9):e12948.

Cohen R, Havlin S, Ben-Avraham D (2003) Efficient immunization strategies for computer networks and populations. *Physical review letters* 91(24):247901.

Feld SL (1991) Why your friends have more friends than you do. *American Journal of Sociology* 96(6):1464–1477.

Godes D, Mayzlin D (2009) Firm-created word-of-mouth communication: Evidence from a field test. *Marketing Science* 28(4):721–739.

Goldenberg J, Han S, Lehmann DR, Hong JW (2009) The role of hubs in the adoption process. *Journal of Marketing* 73(2):1–13.

Hansen LP (1982) Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the econometric society* 1029–1054.

Iyengar R, Van den Bulte C, Valente TW (2011) Opinion leadership and social contagion in new product diffusion. *Marketing Science* 30(2):195–212.

Kim DA, Hwong AR, Stafford D, Hughes DA, O'Malley AJ, Fowler JH, Christakis NA (2015) Social network targeting to maximise population behaviour change: a cluster randomised controlled trial. *The Lancet* 386(9989):145–153.

King CW, Summers JO (1970) Overlap of opinion leadership across consumer product categories. *Journal of Marketing Research* 43–50.

Kumar V, Krackhardt D, Feld S (2018) Network interventions based on inversity: Leveraging the friendship paradox in unknown network structures. *Working Paper, Yale University* .

Libai B, Muller E, Peres R (2013) Decomposing the value of word-of-mouth seeding programs: Acceleration versus expansion. *Journal of Marketing Research* 50(2):161–176.

Stark TH (2018) Collecting social network data. *The Palgrave Handbook of Survey Research*, 241–254 (Springer).

Stephen AT, Lehmann DR (2016) How word-of-mouth transmission encouragement affects consumers' transmission decisions, receiver selection, and diffusion speed. *International Journal of Research in Marketing* 33(4):755–766.

Tucker C (2008) Identifying formal and informal influence in technology adoption with network externalities. *Management Science* 54(12):2024–2038.

Van den Bulte C, Lilien GL (2001) Medical innovation revisited: Social contagion versus marketing effort. *American journal of sociology* 106(5):1409–1435.

Yoganarasimhan H (2012) Impact of social network structure on content propagation: A study using youtube data. *Quantitative Marketing and Economics* 10(1):111–150.

# Electronic Companion Supplement

## EC.1. Mathematical Notation

In Table EC.1 below, we define the terms used in networks. These terms are helpful when we define network properties and in the moment conditions.

**Table EC.1 Table of Notation**

| Characteristic | Description | Definition |
|---|---|---|
| Nodes Degree | Number of connections (edges) of $i$ | $D_i$ |
| Edge | Connection between nodes $i$ and $j$ | $e_{ij} \in \{0,1\}$ |
| Adjacency (Edge) Matrix | Connection between nodes $i$ and $j$ | $\mathbf{E}, E_{i,j} \in \{0,1\}$ |
| Node Set | Set of all $N$ nodes in Network | $\mathcal{V} = \{1,2,\ldots,N\}$ |
| Edge Set | Set of all edges in Network | $E = \{(i,j) : e_{ij} = 1\}$ |
| Network Edge Count | Number of undirected connections | $e = \sum_{i\in\mathcal{V},j>i} e_{ij}$ |
| Seeds | Set of all nodes chosen as seeds | $\mathcal{S}$ |
| Adopters | Set of all nodes which have adopted | $\mathcal{A}$ |
| Reachable Set | Nodes with adoption status $s \in \{A, NA\}$ reachable from $i$ in $k$ steps | $E_i^s(k)$ |
| Proportion of adopting neighbors | Fraction of adopting nodes among those reachable from node $i$ in $k$ steps | $z_i(k) = \dfrac{\lvert E_i^A(k)\rvert}{\lvert E_i^A(k)\rvert + \lvert E_i^{NA}(k)\rvert}$ |
| Vector of above | Vector of adopting proportion of neighbors for each node | $z(k) = [z_1(k),\ldots,z_N(k)]$ |
| Minimum Distance | Distance of Shortest Path between $i$ and $j$ | $\delta_{ij} = \min_k s.t. E^k_{(i,j)} > 0$ |

## EC.2. Identification of Leader Fixed Effect

We demonstrate below that the WOM communication probability for leaders $q_L$ is separately identified from the word of mouth communication probability $q$ for non-leaders. While the argument itself is non-parametric and does not rely on a specific functional form, our demonstration model uses a simple parametric representation consistent with the paper. For this argument, we choose to add a leader fixed effect to the simplest model (Model 1) from the paper.

Suppose we had only static adoption data, we would not be able to identify the fixed effect. However, (i) the availability of time series aggregate adoption data and (ii) the presence of multiple networks allows us to identify the fixed "leader" effect.

First, we note that using only the final adoption levels *will not allow* leader fixed effect $q_L$ to be identified separately from just overall propensity to communicate $q$. Increases in each of these parameters will result in higher final adoption levels in a network. It is straightforward to see that a relatively low level of $q$ in conjunction with a high level of $q_L$ might result in the same adoption level as a high level of $q$ and a low level of $q_L$.

However, the curvature of the adoption trajectory over time provides variation that permits identification of the leader effect $q_L$ separately from $q$. Intuitively, if $q_L$ is higher, the adoption trajectory shows a steeper increase in the earlier periods, since only leaders are communicating initially, and only in subsequent periods do non-leaders communicate. Thus, the proportion of communication attributable to leaders is highest at the beginning and decreasing over time. Thus, the impact of a higher $q_L$ will be greatest in earlier periods as opposed to later periods. In contrast, the impact of a higher $q$ will be lower in the initial periods, since few non-leaders are informed, and it has proportionally greater impact on adoption in later periods.

While the above argument is non-parametric and does not rely on specific functional forms for identification, for the purpose of illustration, we use a parametric model below.

**Simplified Model**

We provide a highly simplified version of the model similar to Model 1 in the paper, for the specific purpose of examining identification and making the required variation transparent. The main features of this model are:

1. A few leader nodes are informed initially (similar to the main model).
2. In each period, each informed node communicates with probability (that depends on the node's leadership status). Thus, non-leaders communicate with probability $q$ and leaders communicate with probability $q_L$ with each of its neighbors. Note that in this simplified model, adoption status *does not* impact communication probability.[9]
3. When nodes are newly informed, they have the ability to adopt a product with probability $\gamma = 0.2$. (We don't have any covariates impacting adoption here, unlike in the main model, and do not require the variation obtainable from these covariates).

The WOM communication probability for node $i$ is specified as:

$$p_i = \begin{cases} q, & \text{if } i \text{ is not a leader} \\ q_L = q + q_\ell, & \text{if } i \text{ is a leader} \end{cases}$$

where $q_\ell$ is the leader fixed effect. Recall that the leader fixed effect is the difference between the WOM communication probabilities of leaders and non-leaders.

We demonstrate in Figure EC.1 precisely the variation that is required for this identification. There are several sources of possible variation in the network data. First, we observe that both adoption trajectories for (a) $q = 0.01, q_\ell = 0.08, q_L = q + q_\ell = 0.09$ (red curve) and (b) $q = 0.13, q_\ell = -0.12, q_L = q + q_\ell = 0.01$ (green curve) end up after $T = 5$ periods at the same overall adoption

[9] Even though this additional variation based on adoption status might prove useful as a separate source of identification, our identification argument does not require it.

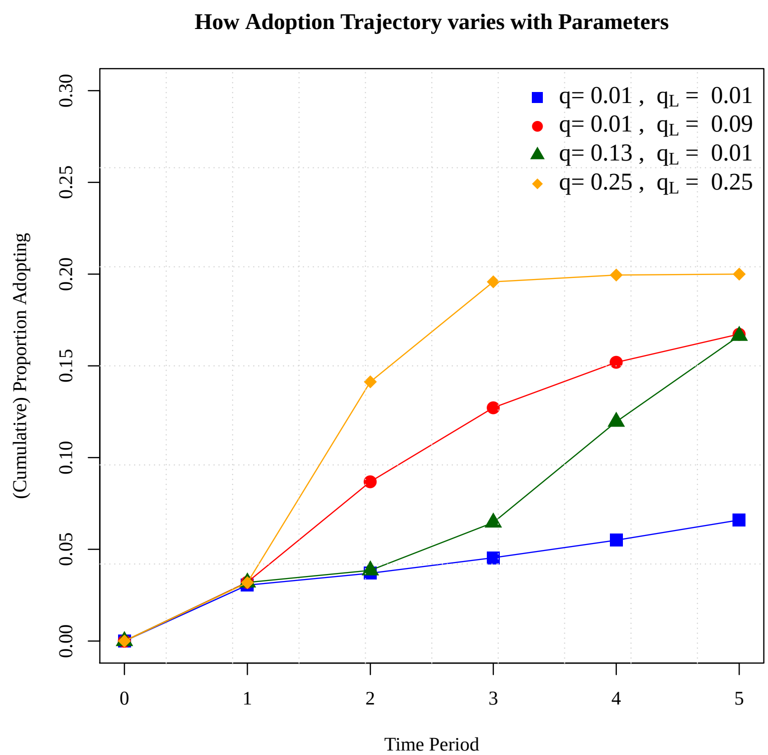



**Figure EC.1** **Identification and Adoption Time Trajectory**

level, i.e. 0.165 or 16.5%. Thus, just having the final adoption levels, it would not be possible to separately identify $q$ and $q_L$.

However, their adoptions differ in their time trajectories. For (a) (red curve), with a higher leader fixed effect $q_L = 0.09$, we see the **early period trajectory is steeper** than the case (b) (green curve). On the other hand, with (b), the later period trajectory is steeper than in (a).

In general, for different combinations of $(q, q_L)$ that *obtain the same level of final overall adoption*, the area under the adoption trajectory curve will be greater for combinations of $(q, q_L)$ with higher levels of $q_L$ and lower levels of $q$.

**Does exclusive seeding by leaders help or hinder identification of leader fixed effect?** There are two reasons why leader seeding (in contrast to random seeding) is helpful to answering our research question.

First, at first glance, it may appear that our context in which the firm exclusively used leaders by the for initial seeding makes it more challenging to separately identify the leader fixed effect. But in fact, our explanation above should clarify that this exclusive use of leaders for initial seeding aids identification of the leader fixed effect and allows us to disentangle $q_L$ and $q$. This is because the exclusive use of leaders for initial seeding guarantees that a higher leader fixed effect will increase the earlier adoption trajectory relative to later. Therefore if the seeding had been random, it would not be feasible to separately identify the effects as one cannot use this identification argument.

Second, leader seeding avoids a specific kind of bias in leader effects. Suppose we only have random seeding, but there are leaders present in the data. If leaders have different (higher or lower) degree on average than others, and if they have differential communication, it would not be possible to identify any leader specific communication effect. For instance, if Leaders have higher degree, the Local friend strategy could result in more leaders on average. The counterfactual results would then be biased to find lower effects for the Local strategy than would be obtained in reality. Due to leader seeding in our data, we can identify and characterize the leader fixed effect (separately from non-leaders), and thus avoid this potential bias.

## EC.3. Model Details and Estimation

First, we detail the estimation of the adoption process, followed by the WOM communication process, and finally detail the block bootstrap to obtain standard errors. We simulated $N_{sim} = 150$ diffusion paths with seeds chosen stochastically corresponding to each seeding level and using each of the seeding strategies. The reported WOM communication parameters are based on the average of the simulated diffusion paths.

### Adoption Process

The adoption parameter vector is $\beta = (\beta_0, \ldots, \beta_6)$. The logistic regression specification for the adoption decision follows from the utility specification. The log likelihood for household $i$ is $l_i(\beta|X_i)$ and for all households in the network is $l(\beta|X)$

$$l(\beta|X) = \sum_{i=1}^{N} l_i(\beta|X_i) = \sum_{i=1}^{N} \log P(y_i = 1|X_i) = \sum_{i=1}^{N} \log \left[ \frac{\exp(\beta X_i)}{1 + \exp(\beta X_i)} \right] \tag{EC.1}$$

(EC.2)

The adoption process is estimated by maximum likelihood estimation.

### WOM Process

Given adoption parameters $\beta$, the WOM process is simulated separately for each village network. We track two states for each household: its information state and its adoption state. The information states are *uninformed* (U) and *informed* (I), whereas the adoption states are *Not-adopted* (NA) and *Adopted (A)*. Both the Informed and Adopted states are absorbing states, during which nodes can communicate with their neighbors.

An *informed household* with adoption status $s \in \{NA, A\}$ (i.e. non-adopting or adopting) will communicate with any of its neighbors in a single time period with probability $p^s(D)$. This is a dynamically evolving process over time, and depends on the informed status of all households in the network. We have formalized these details further below using additional notation. Let $p^s(D)$ be the probability that an *informed household* with adoption status at the beginning of time $t$ $s^j(t) \in \{U, NA, A\}$ (i.e. uninformed, non-adopting or adopting) of degree $D$ will communicate with any of its neighbors *in a single time period.* Uninformed households do not communicate. During time period $t$, an uninformed household $i$ becomes informed if it receives a communication from any of its network neighbors $\mathcal{N}_i$. This event happens with probability $p_{it} = 1 - \Pi_{j \in \mathcal{N}_i}(1 - p^{s_j(t)}(D_j))$.

The WOM process for each of the $N_{sim}$ simulations begins with Step (0) and then proceeds through Steps (1)-(3) for each time period.

(0) Each household (node) in the network is initially in an uninformed (U) information state. In initial period $t = 0$, the seed nodes are chosen in each network based on the seeding strategy. In the actual data, the seed nodes in each village were chosen based on the opinion leadership criterion. In the counterfactual scenarios, seed nodes are chosen based on an alternative strategy (Random, Local Friend etc.). In all cases, the information state of the seed nodes changes from Uninformed (U) $\longrightarrow$ Informed (I).

The following process (1) – (3) process then takes place in each period $t \in \{1, 2, \ldots, T_v\}$ for village $v$.[10]

(1) Each household that has become informed decides whether to adopt.

(2) Then, an informed household can probabilistically communicate about the microfinance product with each of its network neighbors. The probability of such communication $p^s(D)$ may depend on both its degree $D$, i.e. the number of neighbors the informed household has, as well as the adoption status $s \in \{A, NA\}$ of the informed household. We separate out the probabilities $p^{NA}(D)$ and $p^A(D)$ as detailed in §3 of the paper.

(3) When this communication takes place, each neighbor receiving information changes its information state from Uninformed (U) $\longrightarrow$ Informed (I). If the neighbor node has already been informed earlier, there is no change in its state.

For each simulation and for each village $v$, we compute 6 cross-sectional moments according to Table EC.2 at the end of $T_v$ periods of simulation, and 3 time series moments. Thus, for the 43 villages with microfinance adoption, we have $N_{moments} = 9 \times 43 = 301$ moments across the villages. We then minimize the MSM objective function $S(\theta)$ detailed in equation (7) from §3 in the $[0, 1]^K$ region to obtain the probability parameter estimates presented in Table 6 in §4 of the paper. For the MSM objective, we start with the initial weight matrix set to the identity matrix to obtain consistent estimates. Since we obtain standard errors through bootstrap, a consistent estimator is all that is needed.

[10] The number of time periods varies across villages in the data, with a mean of 6.5 and SD of 1.83.

**Standard Errors with Bootstrap Estimation**

We obtain standard errors for the communication probability parameters using a bootstrap procedure detailed below. First, we obtain $N_R = 5,000$ draws using a random grid for the communication probability vector $\theta = (q_0^{NA}, q_1^{NA}, q_0^A, q_1^A) \in [0,1]^4$. The parameter is characterized appropriately based on the model specification.

We proceed through Steps (a) – (c) below for each of the $N_{sim}$ draws to obtain moments for each village $v$.

(a): We choose seeds corresponding to the Leader strategy used in the data.

(b): We compute the simulated WOM Process detailed above for $T_v$ periods for each draw of the parameter vector $\theta$.

(c): We use the cross-section and time series adoption status data to compute the moments detailed in Table EC.2 separately for each village.

Compute $B = 10,000$ bootstrap estimates using the moments obtained from the samples above. For $b = 1, 2, \ldots, B$ do Steps (d) – (f) below.

(d): Resample with replacement from moments from the set of villages showing microfinance activity.

(e): Compute the objective function with the resampled moments at each of the $N_R$ points evaluated above.

(f): Choose the parameter vector with the minimum objective as the estimate $\beta^{(b)}$ to be used in the bootstrap.

The distribution of $\beta^{(b)}$, with $b = 1, 2, \ldots, B$ provides the bootstrap estimate distribution for computing standard errors.

**Moment Conditions for Estimation**

In this section, we describe the rationales for the moments listed in Table EC.2 that we use in our estimation. The required mathematical notation is defined in §EC.1.

In general, all moments are informative in the estimation of all parameters. However, the connections between some moments and parameters are more intuitive. The time series moments, and more generally the temporal trajectory are especially important for identification when there are differential effects for leaders. We describe the moments and the obvious associated links with parameters below.

First, we detail the cross-sectional moments MC1 to MC6. (MC1) is the proportion of seeds that have adopted. Since the seeds are guaranteed to be informed outside the WOM process, this allows us to estimate the parameters impacting adoption probability without relying on the communication process. In contrast, (MC2) is the proportion of households with no adopting

**Table EC.2 List of Moments.**

| Symbol | Description | Definition |
|---|---|---|
| **MC1** | Proportion of seeds adopting | $\frac{\lvert\mathcal{S}\cap\mathcal{A}\rvert}{\lvert\mathcal{S}\rvert}$ |
| **MC2** | Proportion of households with no adopting neighbors who have adopted | $\frac{\sum_{i\in\mathcal{A}}\mathbf{I}\left[\mathcal{N}_i\cap\mathcal{A}=\phi\right]}{\sum_{i\in\mathcal{V}}\mathbf{I}\left[\mathcal{N}_i\cap\mathcal{A}=\phi\right]}$ |
| **MC3** | Proportion of neighbors of adopting seeds who have adopted | $\frac{\bigcup_{j\in\mathcal{S}\cap\mathcal{A}}\lvert\mathcal{N}_j\cap\mathcal{A}\rvert}{\bigcup_{j\in\mathcal{S}\cap\mathcal{A}}\lvert\mathcal{N}_j\rvert}$ |
| **MC4** | Proportion of neighbors of non-adopting seeds who have adopted | $\frac{\bigcup_{j\in\mathcal{S}\cap\mathcal{V}\setminus\mathcal{A}}\lvert\mathcal{N}_j\cap\mathcal{A}\rvert}{\bigcup_{j\in\mathcal{S}\cap\mathcal{V}\setminus\mathcal{A}}\lvert\mathcal{N}_j\rvert}$ |
| **MC5** | Covariance between a household's adoption and average adoption of their first degree neighbors | $cov\left(y,z(1)\right)$ |
| **MC6** | Covariance between a household's adoption and average adoption of their second degree neighbors | $cov\left(y,z(2)\right)$ |
| **MT$\tau$** | Cumulative adoption upto time $\tau$ (Time series moment) | $y_\tau=\frac{1}{N}\sum_{j=1}^{N}y_{j\tau}$ |

neighbors who adopt, which allows us to match a non-adopter's communication likelihood, because such an adopting household could only have received information from neighbors, all of whom are non-adopters.

(MC3) is the proportion of neighbors of adopting seeds who have adopted. This moment most closely connects to the WOM probability of adopters, since the neighbors of seeds have a high probability of receiving information from the seeds. With (MC4), the proportion of nodes that are neighbors of non-adopting seeds who adopt. The focus here is primarily on parameters $q_0^{NA}$ and $q_1^{NA}$. With low probability, it becomes less likely that neighbors of non-adopting seeds would adopt (all else being equal).

(MC5) and (MC6) captures the relationship between adoption by a focal household and its first and second degree neighbors. This is particularly important in networks where there is a significant region (or sub-network) that is uninformed. In such regions of the network, both a focal node and its neighbors will have zero adoption, which results in a perfect correlation. Observe that in such a case, (MC2) and (MC4) are not informative since the moment will have values exactly zero for such sub-networks. Thus (MC5) and (MC6) can also be viewed as characterizing the limits of the WOM process.

Overall, we need to have moments that match global network-level measures, e.g. (MC1) that focuses on overall adoption. It is also critically important to incorporate moments that match local network structure, allowing these connections to have a strong impact on the adoption process, which is what distinguishes the network approach from the Bass model.

The time series moments (MT$\tau$) matches the cumulative overall adoption in each time period $\tau$ period within each village. This is the typical data used in estimation of aggregate Bass-like diffusion models. These moments helps us to estimate the time-path of the diffusion process. In each period of the model, based on the network structure and the diffusion of the information process, we have different number of households which potentially become informed and therefore have the opportunity to make adoption choices.

We detail the sensitivity of parameter estimates to moments using the methodology of Andrews et al. (2017) in §EC.5.7.

### EC.3.1. Model Fit

**Additional Model Fit Metrics** We next evaluate the fit of these models below using 3 additional measures. The metrics used for fit are detailed below:

1. First, we regress the actual adoption rate during each time period in the data (as dependent variable) against the simulated adoption rate obtained from the model, similar to what Banerjee et al. (2013) present in Table 2 of their paper. The intercept terms are found to be not significant, and the coefficient of interest across all models indicate that the model is able to capture and characterize the essential dynamics of the process. If the coefficient of simulated adoption is close to 1, that would indicate a good fit.
2. Next, we examine typical fit measure like **RMSE** (root mean squared error) and **MAPE** (Mean Absolute Percent / Proportion Error). Lower values of these measures indicate better fit.

We find that the model fit is consistent with the original paper for in-sample fit (see Table 2 of Banerjee et al. (2013)). We then examine out of sample fit by estimating our preferred models using 85% of the villages, and holding the remaining 15% of the sample as holdout. We find that the out of sample fit is not significantly worse than in sample fit, indicating the models do not suffer from an obvious overfitting problem. Banerjee et al. (2013) do not provide *out of sample fit* in their paper.

Table EC.3 provides the in-sample and out-of-sample fit for our preferred models. We note that the coefficients on simulated adoption for both in-sample and out-of-sample are between 0.87 and 0.89. The RMSE and MAPE measures are similar for both of our chosen models, and it is useful to verify that the out-of-sample fit is not much worse than in-sample fit. If out-of-sample were indeed much worse, then we should be concerned about the model overfitting the data.

Table EC.4 provides results for all the model specifications. We find that across the specifications, the intercept term is not significant and that the models seem to be fairly similar in terms of their fit.

## EC.4. Counterfactuals

We detail first the implementation of each of the strategies, and then performance of the strategies under different models and at different seeding levels.

**Table EC.3 Main Models: In Sample and Out of Sample Model Fit Measures**

| | In Sample Fit | | Out of Sample Fit | |
|---|---|---|---|---|
| | $\mathbf{M_2^{E=1}}$ | $\mathbf{M_2^{E=0}}$ | $\mathbf{M_2^{E=1}}$ | $\mathbf{M_2^{E=0}}$ |
| Intercept | 0.002 | 0.000 | -0.002 | -0.001 |
| | (0.02) | (0.02) | (0.02) | (0.02) |
| Simulated Adoption | 0.874 | 0.89 | 0.875 | 0.87 |
| | (0.097) | (0.098) | (0.096) | (0.1) |
| RMSE | 0.067 | 0.067 | 0.069 | 0.069 |
| MAPE (×100%) | 0.379 | 0.372 | 0.395 | 0.406 |

**Table EC.4 Additional Fit Measures for Models**

| | $\mathbf{M_1^{E=0}}$ | $\mathbf{M_2^{E=0}}$ | $\mathbf{M_3^{E=0}}$ | $\mathbf{M_4^{E=0}}$ | $\mathbf{M_1^{E=1}}$ | $\mathbf{M_2^{E=1}}$ | $\mathbf{M_3^{E=1}}$ | $\mathbf{M_4^{E=1}}$ |
|---|---|---|---|---|---|---|---|---|
| Intercept | 0.002 | 0.002 | -0.003 | -0.002 | 0.003 | 0.000 | -0.004 | 0.001 |
| | (0.021) | (0.02) | (0.02) | (0.021) | (0.019) | (0.02) | (0.02) | (0.021) |
| Simulated Adoption | 0.877 | 0.874 | 0.901 | 0.899 | 0.867 | 0.89 | 0.904 | 0.876 |
| | (0.103) | (0.097) | (0.100) | (0.105) | (0.096) | (0.098) | (0.099) | (0.104) |
| RMSE | 0.068 | 0.067 | 0.067 | 0.068 | 0.067 | 0.067 | 0.067 | 0.068 |
| MAPE (×100%) | 0.385 | 0.379 | 0.375 | 0.384 | 0.379 | 0.372 | 0.377 | 0.393 |

### EC.4.1. Seeding Strategy Implementation

Table EC.5 provides specific implementation details for each of the seeding strategies we consider.

### EC.4.2. Comparison of Strategies

Figure EC.2, shows the performance of the strategies pairwise, where performance is measured by the proportion of informed households in each counterfactual strategy evaluation.

Next, we examine the consistency of relative performance of the various seeding strategies across villages. **??** provides an overall comparison of the 4 strategies with the adoption levels of Leader, Local Friend and Hybrid strategies plotted against one another. We find that both Local Friend and Friend of Leader consistently perform better on adoption relative to Random as all villages fall above the diagonal. In contrast, while Leader is better than Random for most villages, it is worse for some villages, as shown by the points that fall above the diagonal in the top-left panel.

Moreover, the villages where the Leader strategy performs especially well are smaller (fewer households). In terms of the hybrid strategies, we find that the weak hybrid Friend of Leader strategy mostly outperforms Leader, but it does not do better than Local Friend overall. The strong hybrid Leader Friend of Leader actually performs worse than the Local Friend and weak hybrid strategy. In many villages, it performs worse than the Leader strategy as well.

### EC.4.3. Leverage Under Different Models

We examine how the number of seeds impacts the performance of different seeding strategies in the counterfactual across the full set of model specifications. We examine seeding at the level of

**Table EC.5 Seeding Strategies and Implementation**

| **Category** | **Strategy** | **Implementation Procedure (for each of $m$ seeds)** | **Information Required** |
|---|---|---|---|
| Random | Random | Select node at random from list as seed. | Randomly sampled subset of list of individuals (or Complete List) |
| Friend | Local Friend | Select node at random from list. Obtain one randomly chosen friend of node as a seed. | Randomly sampled subset of list of individuals + Obtain random friend |
| Leader | (Firm's) Leader | Select node from list of leaders indicated by firm | List of Leaders (where leadership is specific to domain) |
| | Like Leader | Select leader node $\ell$ at random. Select the non-leader node most similar to $\ell$ in terms of network properties‡. | List of leaders + Entire Social Network (Adjacency Matrix $\mathbf{E}$) |
| Hybrid | Friend of Leader (Weak Hybrid) | Select a random **leader** from list of leaders. Obtain one randomly chosen friend of this leader as a seed. | List of leaders + Obtain **random** friend |
| | Leader Friend of Leader (Strong Hybrid) | Select a random **leader** from list of leaders. Obtain one randomly chosen friend who is also a **leader** to be seed. | List of leaders + List of **leader** friends of each leader |
| Network Information Strategies | Top Degree | Select a **node** randomly from list of top degree nodes (We specify this as the top 15% most highly connected nodes. | Degree of each node in the network. |
| | Top Diffusion | Select a **node** randomly from list of top diffusion centrality nodes. Diffusion Centrality is defined as $DC = [\sum_{t=1}^{\tau} q^t \mathbf{E^t}] \cdot \mathbf{1}$ where $\mathbf{E}$ is the adjacency matrix and $\mathbf{1}$ is the column vector of 1s. As suggested in Banerjee et al. (2013), we set $q = \frac{1}{\lambda_1}$ where $\lambda_1$ is the greatest eigenvalue of the adjacency matrix. | Full Adjacency matrix $\mathbf{E}$ (who is connected to whom) and the number of periods for diffusion $\tau$. |

‡ : Similarity between nodes in network position could be implemented using the following centrality metrics (among others): degree, eigenvector, Bonancich power centrality

0.5%, 1%, and 5% to understand how the level of seeding affects relative benefits of our friendship paradox strategies. The results for different seeding levels are detailed in Table EC.6.

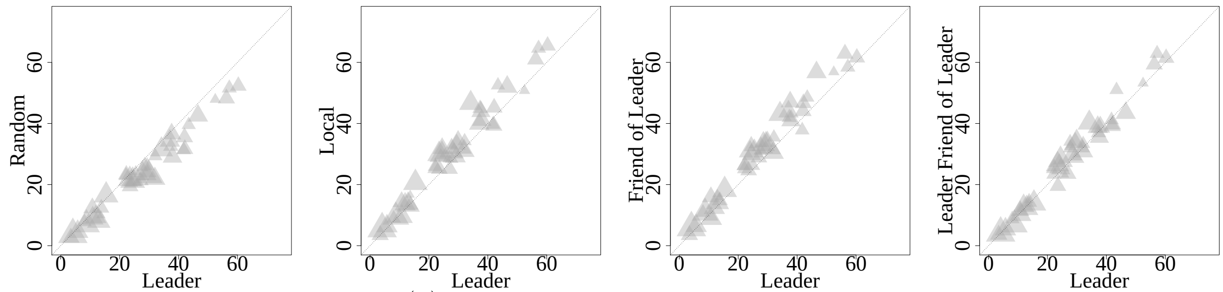
A few observations are relevant here:

(a) The (firm's) leader strategy typically (but not always) outperforms the random node strategy for any combination of model / (#seeds)

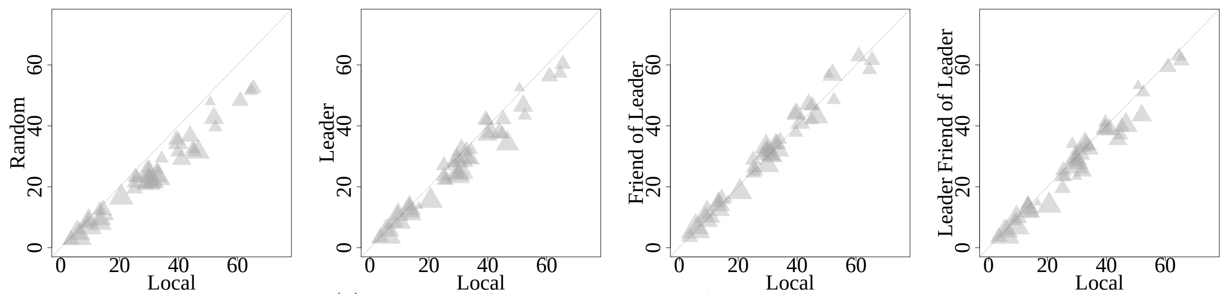
(b) The friendship paradox based Local strategy achieves higher performance (leverage) than the firm's leader strategy under all of the model specifications.

**Figure EC.2 Comparison of Strategies across Villages (1% of Households Seeded). (% Informed Households)**

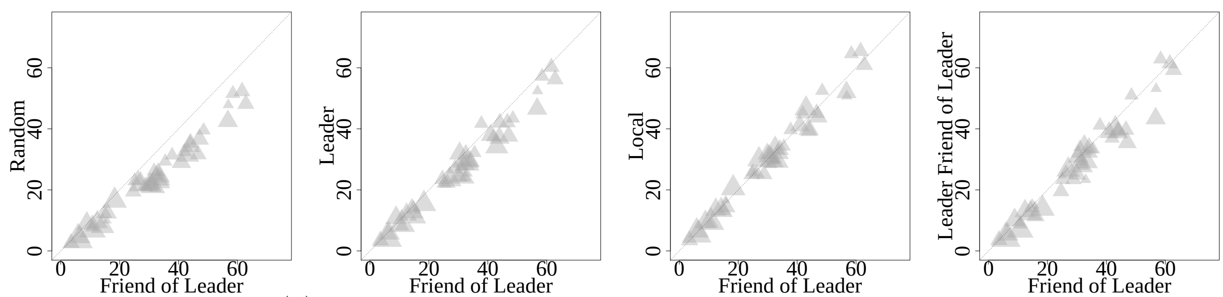
(a) **(Firm's) Leader versus Other Strategies**



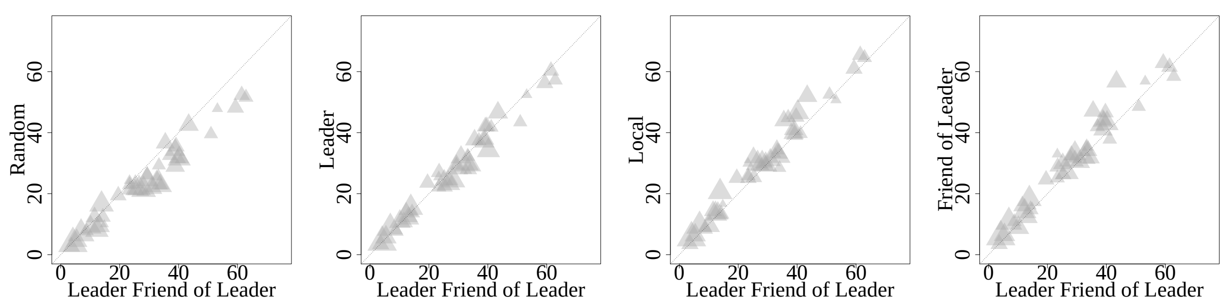

(b) **Local versus Other Strategies**




(c) **Friend of Leader versus Other Strategies**




(d) **Leader Friend of Leader versus Other Strategies**




*Note:* Each data point triangle is a village network in all panels. The size of the shape is proportional to the size of the village (number of households). Darker colors indicate overlap between villages.

(c) The weak hybrid Friend of Leader seeding strategy achieves better performance than Local strategy in most model specifications. However, the strong hybrid Leader Friend of Leader strategy seems to consistently underperform the (firm's) leader strategy.

**Table EC.6 Leverage for Counterfacual Strategies**

(a) Seeding at 5% of number of nodes

| | $\mathbf{M}_1^{E=0}$ | $\mathbf{M}_2^{E=0}$ | $\mathbf{M}_3^{E=0}$ | $\mathbf{M}_4^{E=0}$ | $\mathbf{M}_1^{E=1}$ | $\mathbf{M}_2^{E=1}$ | $\mathbf{M}_3^{E=1}$ | $\mathbf{M}_4^{E=1}$ | $\mathbf{M}_B^{E=0}$ | $\mathbf{M}_B^{E=1}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Local | 1.05 | 1.05 | 1.04 | 1.04 | 1.05 | 1.05 | 1.04 | 1.04 | 1.04 | 1.04 |
| (Firm's) Leader | 1.03 | 1.03 | 1.00 | 1.00 | 1.02 | 1.03 | 1.00 | 1.00 | 1.02 | 1.03 |
| Like Leader | 1.02 | 1.03 | 1.02 | 1.02 | 1.02 | 1.03 | 1.02 | 1.03 | 1.02 | 1.02 |
| Friend of Leader | 1.05 | 1.05 | 1.03 | 1.04 | 1.04 | 1.05 | 1.04 | 1.04 | 1.03 | 1.04 |
| Leader Friend of Leader | 1.03 | 1.02 | 1.00 | 1.01 | 1.02 | 1.02 | 1.00 | 1.01 | 1.02 | 1.02 |
| Top Degree | 1.08 | 1.07 | 1.05 | 1.06 | 1.08 | 1.09 | 1.05 | 1.06 | 1.05 | 1.06 |
| Top Diffusion | 1.07 | 1.07 | 1.04 | 1.04 | 1.06 | 1.07 | 1.04 | 1.05 | 1.05 | 1.06 |

(b) Seeding at 1% of number of nodes

| | $\mathbf{M}_1^{E=0}$ | $\mathbf{M}_2^{E=0}$ | $\mathbf{M}_3^{E=0}$ | $\mathbf{M}_4^{E=0}$ | $\mathbf{M}_1^{E=1}$ | $\mathbf{M}_2^{E=1}$ | $\mathbf{M}_3^{E=1}$ | $\mathbf{M}_4^{E=1}$ | $\mathbf{M}_B^{E=0}$ | $\mathbf{M}_B^{E=1}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Local | 1.21 | 1.18 | 1.16 | 1.15 | 1.22 | 1.24 | 1.13 | 1.20 | 1.18 | 1.20 |
| (Firm's) Leader | 1.14 | 1.08 | 0.99 | 1.01 | 1.12 | 1.15 | 0.99 | 1.02 | 1.12 | 1.13 |
| Like Leader | 1.13 | 1.10 | 1.10 | 1.09 | 1.12 | 1.13 | 1.07 | 1.10 | 1.11 | 1.11 |
| Friend of Leader | 1.21 | 1.20 | 1.16 | 1.17 | 1.21 | 1.24 | 1.15 | 1.17 | 1.19 | 1.21 |
| Leader Friend of Leader | 1.16 | 1.13 | 1.03 | 1.03 | 1.14 | 1.17 | 1.01 | 1.07 | 1.11 | 1.12 |
| Top Degree | 1.44 | 1.40 | 1.32 | 1.30 | 1.39 | 1.45 | 1.29 | 1.34 | 1.32 | 1.37 |
| Top Diffusion | 1.42 | 1.40 | 1.31 | 1.31 | 1.40 | 1.43 | 1.29 | 1.32 | 1.30 | 1.36 |

(c) Seeding at 0.5% of number of nodes

| | $\mathbf{M}_1^{E=0}$ | $\mathbf{M}_2^{E=0}$ | $\mathbf{M}_3^{E=0}$ | $\mathbf{M}_4^{E=0}$ | $\mathbf{M}_1^{E=1}$ | $\mathbf{M}_2^{E=1}$ | $\mathbf{M}_3^{E=1}$ | $\mathbf{M}_4^{E=1}$ | $\mathbf{M}_B^{E=0}$ | $\mathbf{M}_B^{E=1}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Local | 1.25 | 1.26 | 1.22 | 1.22 | 1.22 | 1.28 | 1.22 | 1.22 | 1.20 | 1.25 |
| (Firm's) Leader | 1.18 | 1.14 | 0.97 | 1.00 | 1.11 | 1.13 | 0.97 | 1.04 | 1.12 | 1.16 |
| Like Leader | 1.19 | 1.12 | 1.08 | 1.14 | 1.13 | 1.13 | 1.12 | 1.10 | 1.12 | 1.15 |
| Friend of Leader | 1.32 | 1.26 | 1.21 | 1.23 | 1.28 | 1.27 | 1.24 | 1.25 | 1.21 | 1.27 |
| Leader Friend of Leader | 1.22 | 1.17 | 1.03 | 1.04 | 1.15 | 1.18 | 1.04 | 1.07 | 1.13 | 1.20 |
| Top Degree | 1.60 | 1.53 | 1.44 | 1.47 | 1.53 | 1.55 | 1.47 | 1.48 | 1.45 | 1.58 |
| Top Diffusion | 1.59 | 1.52 | 1.44 | 1.47 | 1.54 | 1.57 | 1.47 | 1.48 | 1.44 | 1.55 |

(d) The "Like Leader" strategy performs very similar to Firm's leader (within 2-3% of the leverage metric).

(e) Leverage for all counterfactual strategies decreases as the number of seeds increases.

## EC.5. Alternative Models and Robustness Checks

We consider different models of WOM communication and seeding to assess if our key claims are sensitive to model specification. Here we consider three models of WOM communication. In §**??**, as a basic benchmark, we consider a single source—advertising type, non-network model where information is not transmitted through the social network, but all households receive information from a central single source (perhaps the firm). Assessing the relative fit of this model with respect to our preferred network based communication model can clarify the importance of modeling information transmission through social networks before even assessing the role of seeding strategies.[11] In §EC.5.2, we consider a broadcast model where in the first period seeds conduct a village-wide meeting in which information about the microfinance program can be broadcast to all those who attend. This model is motivated by a meeting process that the firm encouraged the seeds to conduct and is described in Banerjee et al. (2013). In §EC.5.4, we consider a process where leaders have a certification impact—their adoption or support may increase persuasion even outside of their networks. This model goes beyond the WOM effects of leaders in allowing differential impact based on source of certification. Finally, in §EC.5.3 we also consider a seeding process where not just the firm chosen seeds are leaders but a random sample from certain select occupations that are considered as leaders. In this process, leaders are present both among seeds and non-seeds, but all seeds are randomly chosen members of these occupations.

In the text of the paper, we had used the best fitting model of adoption (Table 5). In §EC.5.5, we report the alternative models of adoption we considered and their relative fit with respect to the chosen model. Finally, we report two sensitivity analysis: a $k$-fold cross-validation for out of sample fit in §EC.5.6 and parameter sensitivity to the different moments used in the estimation §EC.5.7.

### EC.5.1. Single Source Model

We consider a single source—advertising type, non-network model here as a null benchmark model. In this model, information is not transmitted through the social network, but all households receive information from a central single source (perhaps the firm). In **each period**, the source transmits information to each household with probability $\theta$—could be thought of as a household seeing an ad. Informed households then have a chance to adopt. The adoption model is identical to the main model in the paper. Informed households do not communicate any information to other household in this single source model.

Observe that this model is parametrized by only one parameter $\theta$. The estimated value $\theta$ on an average set of villages is $\hat{\theta} = 0.46$. Using this model, we evaluate the in-sample and out of sample fit (using 15-20% of the villages as a hold out sample). Similar to Banerjee et al. (2013), we regress the real adoption data on the simulated adoption trajectory derived from the model.

[11] We thank a reviewer for suggesting this benchmark

**Table EC.7 Null Single Source Model: Fit Measures**

| | *Dependent variable:* | |
|---|---|---|
| | In Sample Villages Adoption Share | Out of Sample Villages Adoption Share |
| Constant | 0.060 | 0.017 |
| | (0.064) | (0.197) |
| Simulated Adoption share | 0.847 | 1.143 |
| | (0.544) | (1.669) |
| RMSE | 0.112 | 0.076 |
| MAPE (×100%) | 0.38 | 0.30 |
| $R^2$ | 0.047 | 0.055 |
| *Note:* | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

There are a few observations:

1. The coefficient of simulated adoption share based on the estimated null model are not statistically significant for either the in sample or out of sample adoption share regressions. Thus, this model does not have any predictive powerin explaining true adoption.
2. For the in-sample fit, the model performs worse than our main models evaluated in the paper.

The finding above is not surprising, since a null model must lead to a concave cumulative adoption curve over time (since there are fewer households that have not been informed over time), whereas a network based model is more consistent with the S-shaped curve for cumulative adoption, similar to the classic model of Bass (1969).

Overall, the empirical evidence of Table EC.7 does not support the single source non-network model, since this model does not capture the primary data patterns of adoption across the village networks.

### EC.5.2. Broadcast during Initial Period

Our model proposed that information about microfinance propagates through word of mouth over the social network. We consider a benchmark (null) model where the information is *broadcast* to households initially in period 0 at a meeting, where the attendance at the meeting is probabilistic. In such a model, information flows directly from a common source to any of the households in the network (subject to their attendance at the meeting), and the *structure of the social network is not relevant for this initial communication.* After this initial broadcast, regular WOM communication occurs through the social network in subsequent periods. As in our main models, informed households have the opportunity to make an adoption decision, whereas non-informed households cannot do so.

We explain why modeling the initial broadcast mechanism would only strengthen our qualitative conclusions about the relative superiority of local friend seeding. In a model with the broadcast mechanism, we should attribute *some* part of the adoption in early periods to that meeting rather than organic household-to-household word of mouth. This implies that the word-of-mouth driven trajectory would be *even lower* in earlier periods, which further implies that the leader fixed effect would be more negative. Taking this logic to the counterfactuals where non-leaders are chosen as seeds, we would therefore see a further increase in relative performance of the local friend seeding strategy and other non-leader strategies when compared to the leader strategy that generates the data.

We now demonstrate this argument by estimating the model that allows for such an initial broadcast and performing the counterfactual. To reiterate, we restate the modeling assumptions.

- In period 0, the leaders invite members of the village to an initial meeting where they explain the microfinance product. For each household, the probability of attending the initial meeting is $\gamma$. We refer to this as the initial broadcast effect.
- In addition to the above, regular word of mouth communication happens through the social network, as we have specified in the main model.
- We consider the case when there are separate leader fixed effects, and the case where this effect is absent.

There are a few points to consider here. First, there is the question of separate identification of an initial broadcast effect from the leader fixed effect. Here, the separation is possible because the broadcast is a one-time initial event, whereas the impact of the leader fixed effect continues beyond the initial period. Thus, if we have 3 or more periods, we can identify both effects. Next, a higher level of leader fixed effect ($q_\ell$) will lead to more friends of leaders being informed (relative to non-friends), and leads to higher adoption among friends of leaders. In contrast a higher value of initial broadcast effect $\gamma$ informs households who are not friends of leaders, leading to higher adoption among that group.

Second, if the probability of attending the initial meeting is very high, $\gamma \approx 1$, then there is little role for the network in communication. The model is then similar to the single source null model of §EC.5.1. More generally, the higher the broadcast effect, the less important are the structure of network connections. Third, in the counterfactual where we choose non-leaders as seeds, we might expect this initial broadcast to be less likely or absent. When leaders have a unique ability to do bring about such a broadcast that non-leaders do not possess, then in the counterfactual, we would set the broadcast parameter $\gamma = 0$. Of course, this assumption *stacks the deck against* any of our proposed strategies, but we include it to show that our strategies still perform better than the leader strategy.

Table EC.8 details the parameter estimates from the initial broadcast model. We find that the results are qualitatively very similar to that of our main model, and quantitatively the relative magnitudes and ordering between the parameters are also the same. For instance, non-adopters communicate less than adopters, and degree is negatively correlated to probability of communicating with a network neighbor. In the model with Leader fixed effects, the leader fixed effect is negative but not significant. We also find the initial broadcast effect, which represents the probability that each household attended the initial meeting to be $\gamma = 0.013$ or $\gamma = 0.079$ depending on whether the model includes leader fixed effects or not.

**Table EC.8 Initial Broadcast Model Estimates**

| Parameter | Symbol | Model Specification: Estimates (Standard Errors) $\mathbf{M_5^{L=0}}$ (No Leader FE) | $\mathbf{M_5^{L=1}}$ (with Leader FE) |
|---|---|---|---|
| Non-adopter lowest degree | $q^{NA}_{min}$ | 0.103<br>(0.062) | 0.073<br>(0.062) |
| Non-adopter highest degree | $q^{NA}_{max}$ | 0.063<br>(0.087) | 0.071<br>(0.089) |
| Adopter lowest degree | $q^{A}_{min}$ | 0.401<br>(0.034) | 0.392<br>(0.090) |
| Adopter highest degree | $q^{A}_{max}$ | 0.314<br>(0.100) | 0.259<br>(0.129) |
| Leader Effect | $q_\ell$ | –<br>– | -0.001<br>(0.080) |
| Initial Broadcast Effect | $\gamma$ | 0.013<br>(0.030) | 0.079<br>(0.030) |
| SMM Objective<br>*J-Statistic* ($\times 10^{-6}$) | | 5.303 | 5.986 |

**Table EC.9 Leverage with Initial Broadcast Model**

| Strategy | *Seeding at:* | No Leader FE 0.50% | No Leader FE 1.00% | No Leader FE 5.00% | With Leader FE 0.50% | With Leader FE 1.00% | With Leader FE 5.00% |
|---|---|---|---|---|---|---|---|
| Local Friend | | 1.26 | 1.14 | 1.08 | 1.21 | 1.10 | 1.03 |
| (Firm's) Leader | | 1.14 | 1.09 | 1.04 | 1.15 | 1.06 | 1.02 |
| Like Leader | | 1.08 | 1.08 | 1.04 | 1.13 | 1.06 | 1.02 |
| Hybrid Strategies: | | | | | | | |
| Friend of Leader (Weak) | | 1.26 | 1.15 | 1.07 | 1.22 | 1.10 | 1.04 |
| Leader Friend of Leader (Strong) | | 1.13 | 1.08 | 1.03 | 1.14 | 1.07 | 1.02 |

Table EC.9 reports how the counterfactual strategies perform relative to random seeding using the estimates from the Initial broadcast model. The ratio (leverage) of 1 indicates that the strategy performs just as well as random.

Further, we observe that:

- Leverage reduces as the seeding proportion increases, similar to the main model.
- At all seeding proportions $(0.5\%, 1\%, 5\%)$, the results show that incorporating the initial broadcast effect does not change the relative performance of the Random, Local and Leader seeding strategies, and this holds *with* or *without* leader fixed effects.

### EC.5.3. Leader Based on Occupation

In our main model, we made the assumption that the set of households with leader fixed effects is the same as the seed set, as all of the seeds were considered leaders in their villages. But as per Banerjee et al. (2013), the microfinance firm chose its seeds based on whether they were in certain “leader” occupations (e.g., teachers, shopkeepers, business owners etc). However, not all households with those occupational characteristics were chosen as seeds.

It may then be reasonable to consider a specification where the seeds are assumed to be randomly sampled from those working these selected occupations. In this case, the leader fixed effect should be associated with all members belonging to these occupations, *whether they were used as seeds or not*.

We estimate such a model and obtain parameter estimates $\theta = (q^{NA}_{min}, q^{NA}_{max}, q^{A}_{min}, q^{A}_{max}, q_l) = (0.074, 0.056, 0.424, 0.344, -0.016)$. We then run counterfactuals under different seeding strategies as before.

The counterfactual results presented in Table EC.10 show that the results are quantitatively similar and qualitatively identical. The results for all the strategies differ from the main results of Table 10 because the parameter estimates used above are different.

**Table EC.10 Leverage for Counterfactual Strategies**

| Strategy | Seeding Level | 0.5% | 1% | 5% |
|---|---|---|---|---|
| Leader by Occupation | | 1.17 | 1.13 | 1.03 |
| Local Friend | | 1.27 | 1.19 | 1.05 |
| Friend of Leader | | 1.27 | 1.24 | 1.05 |
| Leader Friend of Leader | | 1.16 | 1.15 | 1.04 |
| Like Leader | | 1.16 | 1.11 | 1.04 |
| Top Degree | | 1.53 | 1.40 | 1.08 |
| Top Diffusion | | 1.53 | 1.37 | 1.07 |

### EC.5.4. Leader Certification Effect

One concern is if the leader seeds chosen by the firm had an additional certification effect beyond their differential WOM communication that we have already modeled as the leder fixed effect. The original study seeded information with leaders who were pre-defined. Thus, there might be a question of whether such an effect may be present in the counterfactual, where seeding is not focused on leaders. Such a *"leader certification effect"* effect posits that *households might be more likely to adopt if they hear through word of mouth that a leader has certified or endorsed the product.*

We detail two arguments below to demonstrate that this concern is unlikely to hold in the counterfactual. First, for this concern to be valid, none of the seeds recommended by the counterfactual strategy should overlap with the "leader characteristics" of seeds chosen in the original study. To the extent, those choices were made based on certain occupations and other characteristics, this is unlikely that our seeding strategies did not have overlap with the chosen occupations.

Second, we quantify the overlap in leaders (which are originally chosen by the firm) across the strategies, and demonstrate that such leaders are chosen even in the counterfactuals, although more under some strategies than others.

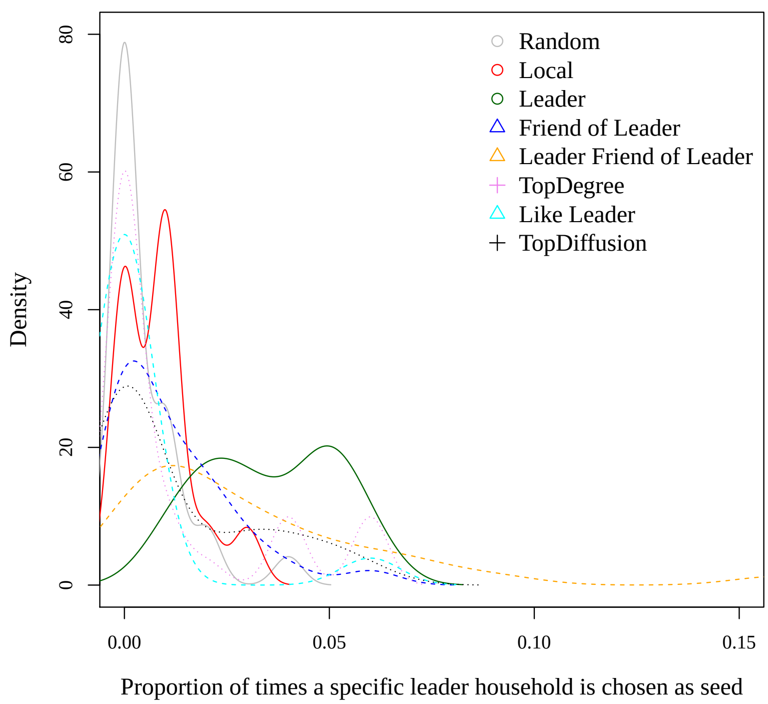



**Figure EC.3 Seeding Overlap (1% seeding)**

We detail the degree of overlap with a density plot in Figure EC.3 that details the probability that any given leader household will be chosen under each of the counterfactual strategies.[12] We focus on seeding at 1%, the results are qualitatively similar at other levels. There are a few noteworthy observations here. First, we do expect that almost all leader households have a higher probability

[12] The main strategies (Random, Local Friend and Leader) are in solid lines, whereas the hybrid strategies are in dashed lines and the network information strategies are dotted line format.

of being chosen as seeds under the Leader strategy, since the seeding is limited to leader households here. Second, as expected, the Random strategy (in gray) has the lowest probabilities of these leader households being selected. Third, observe that leader households have a non-zero probability of being chosen under each of the counterfactual strategies, notably the Local Friend strategy. Finally, each of the other (non-random) counterfactual strategies have a higher probability of leader households being selected compared to the random strategy. More specifically, observe that the comparison of interest is the difference between the Leader strategy (in green) and the Local Friend strategy (in red).

Overall, we believe that above result indicates that our counterfactual outcomes are unlikely to be biased due to the potential for some *unobserved leader certification* ability in the chosen seeds.

### EC.5.5. Adoption Model

The adoption model in the main paper was chosen based on model fit across models that incorporated household data, as well as network characteristics of households derived from the social network within each village. We detail the results in Table EC.11. The results suggest that the household characteristics including number of rooms and beds and indicator for electricity are informative for adoption, and produce a better model fit, as measured by the Akaike Information Criterion (AIC). The AIC is a well regarded measure of fit that is commonly used since it balances model complexity with a likelihood based model fit, unlike measures like likelihood and Pseudo $R^2$.

Specifically, as we add more predictors (electricity and latrine) in moving from model (1) to model (2), the fit as measured by AIC increases. Similarly in adding Rooms and Beds per capita to obtain model (3), the fit improves. However, we find that the household's home ownership and roof type are not significant predictors of microfinance adoption and the models (4) and (5) that include these variables are worse in terms of AIC than model (3).

**Table EC.11 Comparison of Adoption Models**

| | *Dependent variable:* | | | | |
|---|---|---|---|---|---|
| | Adoption | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Constant | $-0.560^{***}$ (0.149) | $-1.292^{***}$ (0.320) | $-1.210^{***}$ (0.322) | $-1.353^{***}$ (0.436) | $-1.664^{***}$ (0.612) |
| Number of Rooms | $-0.196^{***}$ (0.058) | $-0.161^{***}$ (0.059) | 0.007 (0.085) | 0.005 (0.085) | 0.004 (0.086) |
| Number of Beds | $-0.108^{*}$ (0.061) | $-0.063$ (0.062) | $-0.283^{**}$ (0.143) | $-0.284^{**}$ (0.143) | $-0.258^{*}$ (0.144) |
| Electricity | | 0.174 (0.122) | 0.156 (0.123) | 0.158 (0.123) | 0.128 (0.125) |
| Latrine | | $0.160^{**}$ (0.080) | $0.179^{**}$ (0.080) | $0.181^{**}$ (0.080) | $0.163^{**}$ (0.082) |
| Rooms Per Capita | | | $-1.023^{***}$ (0.392) | $-1.015^{***}$ (0.393) | $-0.989^{**}$ (0.394) |
| Beds Per Capita | | | $1.147^{*}$ (0.656) | $1.153^{*}$ (0.655) | 1.075 (0.667) |
| Own Home | | | | 0.147 (0.299) | 0.142 (0.304) |
| Roof Type 1 | | | | | 0.729 (0.736) |
| Roof Type 2 | | | | | 0.507 (0.442) |
| Roof Type 3 | | | | | 0.330 (0.443) |
| Roof Type 4 | | | | | 0.397 (0.453) |
| Roof Type 5 | | | | | 0.150 (0.482) |
| Roof Type 6 | | | | | 1.974 (1.561) |
| Roof Type 7 | | | | | 2.109 (1.497) |
| Roof Type 8 | | | | | $-11.716$ (375.669) |
| $PseudoR^2$ | 0.02 | 0.026 | 0.032 | 0.032 | 0.037 |
| Observations | 1,140 | 1,140 | 1,140 | 1,140 | 1,140 |
| Log Likelihood | -610.256 | -606.784 | -603.093 | -602.970 | -599.684 |
| Akaike Inf. Crit. | 1,226.513 | 1,223.567 | 1,220.185 | 1,221.940 | 1,231.368 |

Note: $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01.

The log-likehood $LL_0$ for a model with just the intercept is -622.878 and Pseudo $R^2 = 1 - \frac{LL}{LL_0}$.

### EC.5.6. Cross validation check

This section presents a $k$-fold cross-validation to assess model fit out of sample. We divide the sample into $k = 10$ parts and perform the out-of-sample assessment by estimating the model on a rotating set of $k$-1 parts and testing out of sample fit on a rotating part.

The results of the $k$-fold cross validation for our main model specification are presented in Table EC.12.[13] In Table EC.12a, we provide the mean and standard deviation obtained from the $k$ different parameter estimates. The standard deviation of the parameter estimates is not very high, indicating that the parameter estimates obtained from each of the folds are quite similar, leading to more confidence about the model predictions.

Table EC.12b presents the in sample and out of sample metrics of model fit.The $R^2$ for the in sample and out of sample predictions based on the model are 0.80 and 0.65 respectively. The corresponding in sample and out of sample correlations are 0.88 and 0.86.

Overall, we find that the $R^2$ and the correlation in trajectories for simulated and true adoption are not only high but also quite similar in sample and out of sample. These findings indicate that the model does not suffer from overfitting and lends further credence to the model estimates.

**Table EC.12 Cross Validation**

(a) Model Estimates

| | Estimate | |
|---|---|---|
| | Mean | SD |
| Intercept | -0.04 | 0.01 |
| Simulated Adoption | 1.38 | 0.06 |

(b) Model Fit

| | In Sample | | Out of Sample | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| $R^2$ | 0.80 | 0.03 | 0.65 | 0.18 |
| Correlation | 0.88 | 0.02 | 0.86 | 0.09 |

[13] In this exercise, we use the same controls as in Banerjee et al. (2013), i.e. number of households in village, membership in self-help groups, membership in savings plan, caste / socioeconomic indicator, and proportion of leader households in village.

### EC.5.7. Parameter Sensitivity to Moments

We have explored a wide variety of alternative model specifications and assumption for both the adoption and word of mouth communication processes. However, an interested reader might be interested in testing robustness to an alternative they have in mind that might be quite different. This issue is explored in detail by Andrews et al. (2017), who provide a unified framework to help make structural (and other) models more transparent so that readers can easily evaluate sensitivity to assumptions. They recommend providing a sensitivity matrix ($\Lambda$) that allows us to evaluate how violations of specific moment conditions can change the model parameter estimates.

One might view the approach of Andrews et al. (2017) as complementary to specifying different models to demonstrate robustness, which we have also done. Their point is that since it is impossible to test *all potential alternative models*, providing the sensitivity matrix allows any interested reader to determine how each of the data moments contribute to parameter estimates. Following Andrews et al. (2017), we report the sensitivity matrix $\Lambda$ in Table EC.13.

Similar to the applications presented in Andrews et al. (2017), we scale the values so that the sensitivity values correspond to a 1% change in each moment condition. These results can be helpful in evaluating the sensitivity of each parameter on each of the moment conditions used in estimation.

First, observe that we use simulated method of moments (SMM), which is a (simulated) version of Generalized Method of Moments (GMM) to obtain the parameter estimates. This allows for the form of the sensitivity matrix

$$\Lambda = -(G'WG)^{-1}G'W$$

where $G$ is the Jacobian corresponding to the moment conditions $g(\theta)$ and $W$ is the weighing matrix used in the GMM estimation. The main result of Andrews et al. (2017) is that the asymptotic bias of local violations of the moment conditions is then given as

$$\mathbf{E}(\tilde{\theta}) = \Lambda \mathbf{E}(g)$$

so that knowing $\Lambda$ allows us to determine how violations of the moment conditions $g$ translate into differences in parameter estimates.

The sensitivity matrix corresponding to the baseline model with the leader fixed effect are detailed in Table EC.13. For each parameter, the table shows the sensitivity of parameter estimates to violations of the moment conditions. We have 9 moments that are each present across all the villages in the data, and in the model the moments are generated from the parameter values. Overall, there are 6 cross sectional moments, and 3 time series moments (period 1, period 2 and period 3).

**Table EC.13 Plug-in Sensitivity $\Lambda$ for Model with Leader FE Effect**

| | *Cross Sectional* | | | | | | *Time Series* | | |
|---|---|---|---|---|---|---|---|---|---|
| Moment # $\rightarrow$ | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 |
| $q^{NA}_{min}$ | 0.009 | 0.016 | 0.004 | 0.004 | 0.005 | 0.003 | 0.001 | 0.022 | 0.011 |
| $q^{A}_{min}$ | 0.003 | 0.005 | 0.001 | 0.001 | 0.002 | 0.002 | 0.000 | 0.006 | 0.003 |
| $q^{NA}_{max}$ | 0.008 | 0.050 | 0.005 | 0.001 | 0.018 | 0.014 | 0.000 | 0.090 | 0.037 |
| $q^{A}_{max}$ | 0.013 | 0.012 | 0.011 | 0.013 | 0.004 | 0.011 | 0.002 | 0.003 | 0.005 |
| $q_{\ell}$ | 0.001 | 0.006 | 0.001 | 0.005 | 0.001 | 0.002 | 0.000 | 0.003 | 0.002 |

To understand the table, consider for example the parameter $q^{NA}_{max}$, which represents the word of mouth communication probability for (high degree) non-adopters. First, observe that the cross-sectional moment this parameter is most sensitive to is moment 2, which is the proportion of households with no adopting neighbors who have adopted. Similarly we find that the parameter is not sensitive at all to moment 7, which is the initial adoption. This is consistent because households adopting in the first period are unlikely to be hearing about it from non-adopters. In the earliest period, only the seeds have a chance to adopt, and the seeds are directly informed by the firm, so word of mouth among non-adopters is unlikely to play a role. In contrast, the adoption in periods 2 and later are relevant to the parameter, especially in contrast to the initial period. If initial period adoption is low (consider the extreme case of zero adoption in the initial period), then adoption in later periods must be driven by word of communication, which then informs both parameters $q^{NA}_{min}$ and $q^{NA}_{max}$.

Similarly, we find that for the leader fixed effect, moment 2 and moment 4 are most important. The time series moment 7 (adoption in period T=1) **does not contribute any information** about the leader fixed effect. This is consistent with our intuition since the leader fixed effect is communication by the leaders to their friends, and thus, the level of period 1 adoption will not be informative of how much communication has occurred through the network by leaders. In contrast, period 2 adoption (moment 8) is important since we would find a greater jump in the early adoption trajectory when the leader fixed effect is greater.

Overall, this method provides transparency in illustrating what variation in the moments is driving the parameter estimates of the model.