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*. But higher-degree nodes may communicate less with neighbors; 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 15-20% more effective than random seeding in increasing adoption. Remarkably, they are also about 5-11% more effective than opinion leader seeding, and are relative more effective when we have fewer seeds.

Latest version at: http://faculty.som.yale.edu/vineetkumar/research/BuzzFriends.pdf

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).

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 (degree or eigenvector or betweenness) 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 who 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 desired 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 (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 has a mathematical foundation and holds independent of network structure, because popular people are always overrepresented when averaging over friends (Feld 1991). This over-representation 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. In fact, we show that network characteristics can impact the effectiveness of these seeding strategies.[1]

The above discussion motivates our research questions below.

(1) Can friendship paradox based seeding strategies improve WOM and adoption relative to random and opinion leader based strategies? By how much?

(2) Can hybrid approaches leveraging the friendship paradox along with an individual's "leader" characteristics lead to higher adoption?

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

(4) How does network structure moderate the relative effectiveness of the seeding 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 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 typically the necessary multi-network data is 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 not possible to identify the effect of seeding choices without multiple diffusion paths across similar networks. Finally, the impact of WOM might be mis-identified in the presence of advertising (Van den Bulte and Lilien 2001).

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

[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.

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.

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.

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, 2 and 3 provide the summary statistics of household characteristics, village social networks and microfinance adoption in the villages. In Table 3, 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. 91% 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 Households

Statistic	Mean	St. Dev.	Min	Max
Number of People in Household	4.484	0.535	3.337	5.832
Rooms	2.308	0.413	0.754	2.939
Beds	0.878	0.455	0.293	2.268
Electricity Indicator	0.915	0.114	0.234	0.982
Latrine Indicator	0.290	0.155	0.020	0.909
Proportion of Leaders (%)	12.552	3.159	7	21

Table 2 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 households. Across villages, the mean degree (connections) of households is around 9, The mean

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

Table 2 Summary Statistics of Village Networks

Statistic	Mean	St. Dev.	Min	Max
Number of Households	212.233	53.536	107	341
Number of Connections	1,031.791	334.113	365	2,015
Degree:				
Mean	9.656	1.642	6.822	13.593
Standard Deviation	7.085	1.321	5.175	11.019
Minimum	1.000	0.000	1	1
Maximum	39.721	13.010	23	90
Mean of Leaders	12.935	2.594	8.880	18.818
Other Network Characteristics:				
Density	0.048	0.013	0.024	0.077
Global Clustering	0.198	0.037	0.129	0.282
Average Path Length (APL)	2.770	0.207	2.432	3.316
Degree Assortativity	−0.078	0.054	−0.260	0.090

Table 3 Summary Statistics of Adoption (%)

Statistic	Mean	St. Dev.	Min	Max
Population	19.382	8.160	8	45
Leader	24.713	12.637	3.571	53.846
Followers	18.677	8.190	7.296	43.713
Electrified Households	19.005	8.384	7.339	45.122
Non-electrified Households	20.716	13.256	0	50
Latrine Households	14.685	9.167	0.000	36.364
Non-latrine Households	21.852	9.893	7.031	51.250

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. Given the much higher than average degree of the opinion leaders, the relative superiority of our proposed friendship paradox based seeding relative to opinion leader seeding is truly an empirical question. 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.

In terms of other network characteristics, we consider four "classic" or most important structural features of networks: (a) density, (b) average path length (APL), (c) clustering, and (d) degree assortativity. Density captures how well connected a network is, and is characterized by the ratio of connections to all possible combinations of household pairs. The average path length (APL) is a measure of reachability, and a lower value of this metric indicates a network in which it is possible to reach any node from another node through a short number of intermediate nodes. This notion is related to the popular idea of "six degrees of separation." Clustering is a measure of transitivity, indicating the propensity for friends of an individual to be mutual friends with each other. Networks that demonstrate high clustering feature close and tightly bound community structures, whereas in networks with low clustering, dyadic relationships are of primary importance. Finally, degree assortativity captures whether households with similar degree are connected to each other. When high (low) degree to other high (low) degree nodes), networks have positive assortativity, when high (low) degree nodes are connected to low (high) degree nodes, there is negative assortativity.[3] The precise definitions of network characteristics are provided in the Supplement (Table EC.2). We will investigate how the relative performance of different seeding strategies varies based on these network characteristics.

The primary objective of study here is the adoption of microfinance by households across the villages, detailed in Table 3. 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 correlation 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. Our model is based on Banerjee et al. (2013), with a key adaptation needed to study our research question related to the friendship paradox. The critical difference is that we allow the probability of WOM from a node to differ by degree and for those identified as "leaders" by the firm. Banerjee et al. (2013) allow the probability of WOM

[3] As a point of comparison, Facebook and Twitter have very low density relative to the village social networks at 2.01×10^{-4} and 8.463×10^{-7} respectively; while their assortativity is more negative than village networks at -0.67 and -0.88 respectively (Kunegis 2013).

to depends on adoption status but not on the number of connections (degree) or for "leaders". It is helpful to understand why we need to extend their model.

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). Our approach based on Friendship Paradox obtains higher degree nodes than average. Thus, we chose a conservative approach, allowing for the idea that which 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 would 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 the 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 with adoption status 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, β 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.

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.

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 λ 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. 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_ℓ to be positive, whereas a negative value would indicate leaders are less effective than others.

Non-linear Impact of Degree

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

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

where q_Q represents the parameter corresponding to the quadratic term.

Estimation

The model estimation proceeds in three steps similar to Banerjee et al. (2013), with some differences as detailed in Supplement §EC.2. Here we provide a high level description of the three steps. For estimation details, see Section EC.2.

Step 1: Adoption Process – We estimate the adoption process parameters β with a logistic regression using the adoption decisions of only the initially seeded individuals based on equation (3).

Step 2: WOM Process – We estimate the WOM process parameters $(q_{min}^{NA}, q_{max}^{NA}, q_{min}^{A}, q_{max}^{A})$ using the method of simulated moments (MSM). Given the data, we use the same set of seven moments

Table 4 List of Moments

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

used in Banerjee et al. (2013) 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 captures the overall adoption over time across a network. This is the only time series moment. The remaining six are cross-sectional moments. The second moment is global, matching overall adoption levels in the network. Moments 3-5 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 6-7 are also local moments in that they capture covariance in adoption between a household and its first and second degree neighbors respectively.[4]

The objective function for the parameter vector θ 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, and 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.

4. Results

Table 5 reports the adoption model estimates. 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 poorer households without access to traditional banking services, but that the poorest households are not the biggest adopters.

[4] We provide precise specification of the moments and the rationale for using them in the supplement (Section EC.3).

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	

We examine a number of different models, summarized in Table 6. 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). However we used optimization algorithms for estimation rather than grid search. 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 6 Summary of WOM Model Components

	$\mathbf{M_1^{E=0}}$	$\mathbf{M_2^{E=0}}$	$\mathbf{M_3^{E=0}}$	$\mathbf{M_4^{E=0}}$	$\mathbf{M_1^{E=1}}$	$\mathbf{M_2^{E=1}}$	$\mathbf{M_3^{E=1}}$	$\mathbf{M_4^{E=1}}$
No Endorsement	✓	✓	✓	✓				
Endorsement					✓	✓	✓	✓
Degree-Independent WOM	✓				✓			
Degree-dependent WOM		✓	✓	✓		✓	✓	✓
Leader Differential WOM			✓	✓			✓	✓
Quadratic Effect: WOM and Degree				✓				✓

Table 7 reports the estimates for the 8 WOM models. We use the SMM J-statistic (a measure of fit) for model selection. 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.[5]

Table 7 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)
Endorsement	λ	–	–	–	–	-0.036	-0.007	-0.021	-0.034
		–	–	–	–	(0.046)	(0.076)	(0.067)	(0.07)
Leader Effect	q_ℓ	–	–	-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)
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_ℓ 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.

[5] The counterfactual performance under *all* of the models are provided in the Supplement of the paper (Section EC.7). We also discuss additional model fit metrics in the Supplement in Section EC.4, evaluating both in-sample and out-of-sample fit for different model specifications.

Finally, we find no evidence of an endorsement or persuasion effect, estimated through parameter λ 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.

5. Counterfactuals

We use counterfactuals to evaluate various seeding strategies based on Friendship, Leadership and Hybrid categories. Specifically, we consider the friendship paradox based Local Friend strategy, Opinion Leader as well as Hybrid strategies that combine the features of sampling on friends along with information on opinion leaders in Table 8. We examine two different hybrid strategies: choosing a random friend of a leader household (weak hybrid) or choosing a random *leader friend of a leader* household (strong hybrid). We use the random seeding strategy as the benchmark. In our villages, each leader is connected to at least one other leader, so this does not result in an empty set. To evaluate whether the impact of seeding is due to the network position or due to the differential impact by individual characteristics of leaders, we seed with leader-like individuals, similar to leaders along 3 dimensions: degree, eigenvector and power centrality (Bonacich 1987).

Further details including informational requirements are provided in Section EC.5.

Table 8 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	Leader	Select node from list of leaders
	Like Leader	Select leader node ℓ at random. Select the non-leader node most similar to ℓ 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.

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 6. 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 in Section 5.1 below.

We evaluate seeding effectiveness in terms of proportion of informed households and adoption generated by the seeding strategies. Table 9 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 21%, while the improvement over Random for Leader is about 12%. We also find that the Hybrid strategy Friend of Leader performs the best with a 23% 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 profitably combined to achieve higher performance. However, we note that using the Local 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.

Table 9 Comparison of Strategies (1% seeding)

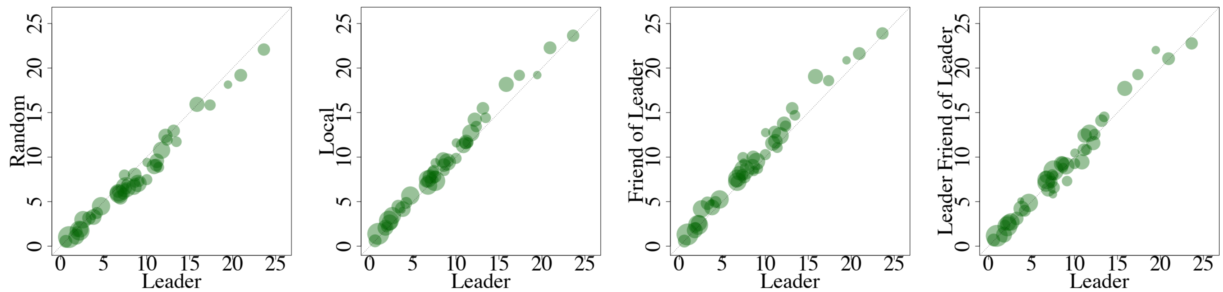
Strategy	**Informed (%)**		**Adopted (%)**		**ΔInformed(%)**	**ΔAdopted(%)**
	Mean	SD	Mean	SD	over Random	over Random
Random	38.410	39.110	7.750	24.440		
Local Friend	45.520	41.020	9.370	26.980	18.520	21.010
Leader	42.700	40.360	8.670	25.880	11.190	11.900
Like Leader	42.970	40.450	8.800	26.120	11.890	13.560
Hybrid Strategies:						
Friend of Leader	46.250	41.190	9.570	27.230	20.410	23.550
Leader Friend of Leader	43.380	40.290	8.890	26.160	12.960	14.770

Next, we examine the consistency of relative performance of the various seeding strategies across villages. Figure 1 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.

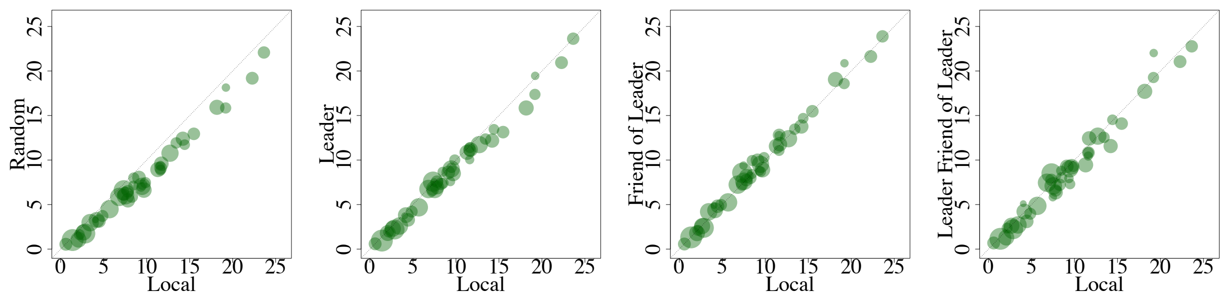
The Local Friend strategy also outperforms the Leader strategy across most villages (88.37%). Moreover, the villages where the Leader strategy performs especially well are smaller (fewer number of households). In terms of the hybrid strategies, we find that the weak hybrid *Friend of Leader*

Figure 1 Comparison of Strategies across Villages (1% of Households Seeded). (% Adopting Households)

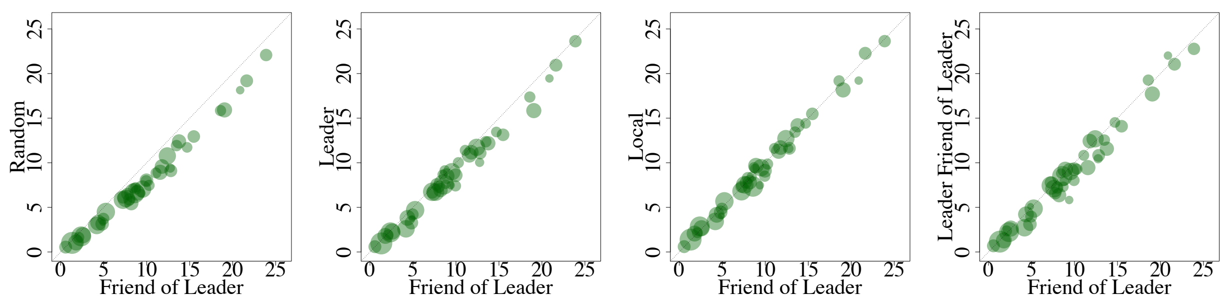
(a) **Leader versus Other Strategies**



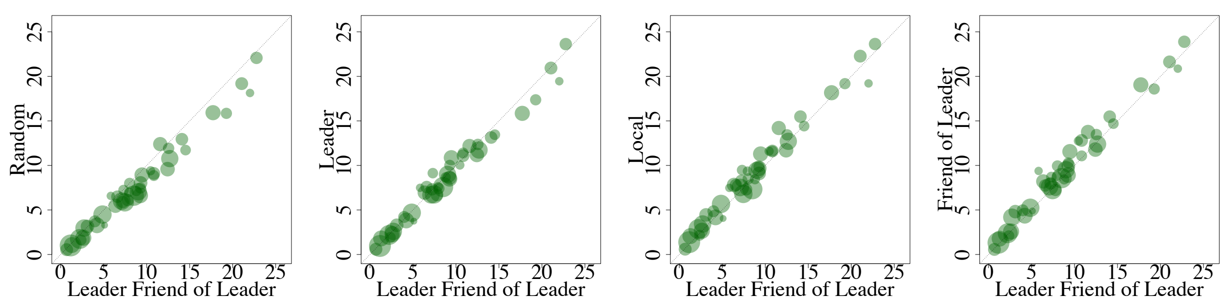
(b) **Local versus Other Strategies**



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



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



Note: Each data point circle 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.

mostly outperforms Leader, it does not do better than Local in general. 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.

Another pertinent issue is noting the region towards the bottom left of the graphs. The villages clustered here display low rates of adoption in general, and thus any improvement made by the proposed strategies in these villages is likely to be especially useful. Note that while the above figure illustrated adoption levels, we detail the communication and proportion of households *informed* in Section EC.5.

Table 10 reports the pairwise comparison in Figure 1 numerically. Local is uniformly better than Random and leads to improved adoption in 100% of the villages. As anticipated from the figure, Leader does worse than random in about 12% of of the villages. The hybrid Friend of Leader strategy is also better than random in all villages, but the hybrid Leader Friend of Leader actually performs worse than random in about 10% 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 friend. The results suggest that it actually reduces 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.

Table 10 **Pairwise Comparison of Strategies (1% seeding)**

	Local Friend Friend	Leader	Like Leader	Friend of Leader	Leader Friend of Leader
Random	100.00	88.37	95.35	100.00	90.70
Local Friend		11.63	20.93	53.49	30.23
Leader			62.79	88.37	67.44
Like Leader				83.72	44.19
Friend of Leader					27.91

Note: % of villages where column strategy achieves higher adoption than row strategy

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 11 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.7.

We define the performance metric as leverage, in terms of how well a proposed seeding strategy s performs relative to Random:

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

Table 11 Leverage for Counterfactual Strategies

Strategy	*Seeding at:*	No Endorsement 0.5%	1%	5%	With Endorsement 0.5%	1%	5%
Local Friend		1.293	1.210	1.116	1.1247	1.236	1.122
Leader		1.160	1.119	1.064	1.123	1.112	1.068
Like Leader		1.181	1.136	1.064	1.120	1.130	1.073
Hybrid Strategies:							
Friend of Leader (Weak)		1.309	1.235	1.119	1.129	1.241	1.126
Leader Friend of Leader (Strong)		1.210	1.148	1.070	1.152	1.160	1.074

The following observations are noteworthy. First, the leader strategy always outperforms the random strategy and the local 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 network position of leaders. Finally, leverage for all strategies decreases as the number of seeds increases.

5.2. How does Network Structure Moderate Seeding Performance?

We evaluate how network characteristics moderate the relative performance (leverage) of the various seeding strategies. We regress leverage as the dependent variable against network summary characteristics from Table 2. This analysis is feasible because we observe diffusion and adoption across several village social networks, in contrast to online scholars typically working with one large social network (e.g., Facebook, Twitter).

Table 12 details the regression estimates. Degree distribution (mean or standard deviation) has no significant effect on performance of seeding strategy. Network density has a positive effect on performance of leader and friend of leader strategies, but not the others.High levels of clustering with strong social communities imply WOM can be easily transmitted within communities but might be more difficult across communities. As such a friend based seeding strategy is unlikely to help improve adoption on highly clustered networks due to similarity in connections. Empirically, we find little effect of clustering on seeding strategy performance, likely due to the limited variance in clustering across village networks, which may be formed by similar social processes.

A low average path length (APL) implies that any node can be reached from any other using a relatively short path. This is typically due to redundancies in paths. With higher levels of redundancy, obtaining a higher degree node is not as valuable for diffusion. However, with low levels of redundancy (or high APL), it becomes more valuable to have well-connected seeds that can reach

Table 12 Performance of Strategies Based on Network Characteristics

	Leverage of Strategy			
	Local	Leader	Friend of Leader	Leader Friend of Leader
Mean Degree	0.029	−0.009	0.019	−0.017
	(0.039)	(0.035)	(0.032)	(0.036)
Std. Dev. of Degree	−0.027	0.019	−0.022	0.046
	(0.044)	(0.038)	(0.036)	(0.040)
Density	1.840	4.219*	3.738*	1.682
	(2.490)	(2.185)	(2.046)	(2.277)
Global Clustering	−0.005	0.756	0.159	0.100
	(0.798)	(0.700)	(0.656)	(0.730)
APL	0.418***	0.290***	0.386***	0.293***
	(0.049)	(0.043)	(0.040)	(0.045)
Assortativity	−0.796	−0.720	−0.327	1.200**
	(0.621)	(0.545)	(0.510)	(0.568)
R^2	0.9875	0.9882	0.9918	0.9879
Residual Std. Error (df = 36)	0.153	0.136	0.125	0.142
F Statistic (df = 7; 36)	422.605***	431.474***	647.992***	423.218***

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

all other nodes through relatively short paths. Consistent with this, we find that an increase in APL improves the efficacy of friendship paradox based seeding strategies.

Local friend should work better when there is negative assortativity, when low degree nodes are connected to high degree nodes. With positive assortativity, low degree nodes are connected to similar low degree nodes. Local friend seeding is likely to yield relatively higher degree nodes when there is negative assortativity, with a star network being an extreme example. Though as expected the signs are negative, they are found not significant. However, the leader friend of leader strategy will benefit from positive assortativity because a leader's leader friend also having high degree benefits the strategy. Indeed the coefficient is positive and significant. Overall, of the 4 characteristics, the most consistent impact is path length in the network.

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 opinion leader seeding. Specifically, we find a 15-20% improvement in both information spread and adoption compared with Random, and about 8% improvement over Leader seeding used by the firm.

We also find that network structure-based strategies can be combined with Leader strategies in hybrid strategies to achieve even higher performance. However, imposing stronger conditions on the hybrid results in poorer performance. Thus, we must balance the somewhat greater informational requirements of the hybrid strategy against higher performance. Overall, the proposed strategies are uniformly better across all the villages with varying network structures, whereas Leader strategies can be worse than Random in a significant number of village networks. We also show that the effectiveness of the seeding strategies depends on network structure, as characterized by summary statistics like average path length and density. A caveat is worth noting: the seeding strategies are stochastic, in the sense that they do not choose pre-determined individuals or households. Thus, any performance guarantees can only be made in expectation.

We find the advantage of both Local Friend and hybrid strategies relative to the random strategy to be inversely related to the proportion of nodes seeded. Thus, when we have few seeds, these strategies become even more advantageous. This might be relevant in cases where the target population is large, and the intervention is somewhat costly, either in monetary terms or in terms of urgency or because of other operational limitations. Finally, we find that structural properties of networks can impact the relative performance of strategies, with higher path length strongly associated with increased performance.

We leave some important issues for future research. First, a promising approach consideres the speed of diffusion and the potential to use seeds nominated by others as "gossipers" as having potential to accelerate diffusion and higher overall adoption (Stephen and Lehmann 2016, Banerjee et al. 2014). An interesting question would be to examine whether friendship could be combined as a hybrid with such approaches (e.g. friend of a gossiper). Second, it would be useful to consider whether seeding approaches proposed here need to be adapted for highly asymmetric networks, where directionality becomes significant (Ben Sliman and Kohli 2018).

References

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.

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.

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* .

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.

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* .

Kunegis J (2013) Konect: the koblenz network collection. *Proceedings of the 22nd International Conference on World Wide Web*, 1343–1350 (ACM).

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. Network 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. Some network characteristics used in our analysis are defined in Table EC.2.

Table EC.1 Network Terminology

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	$\mathcal{V} = \{1,2,\dots,N\}$
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	$z(k) = [z_1(k),\dots,z_N(k)]$
Minimum Distance	Distance of Shortest Path between i and j	$\delta_{ij} = \min_k s.t. E^k_{(i,j)} > 0$

EC.2. Estimation

First, we detail the estimation of the adoption process, followed by the WOM communication process, and finally detail bootstrapping to obtain standard errors. We simulated $N_{sim} = 150$ diffusion paths based on $N_S = 1$ seed chosen 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 $\beta_0,\dots,\beta_6$. The logistic regression specification for the adoption decision follows from the utility specification. The log likelihood for household i is $l_i(\theta|X_i)$ and for all households in the network is $l(\beta|X)$

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

Table EC.2 Network Characteristics

Characteristic	Description	Definition
Density	Ratio of the number of edges to the number of possible edges	$\frac{2\|E\|}{N(N-1)}$
Clustering	Ratio of the number of closed triads (triangles) to the number of possible triads	$\frac{\sum_{i\neq j\neq k} e_{ij}e_{jk}}{\sum_{i\neq j\neq k} 1}$
Average Path Length	Average of Minimum Distances across all pairs of nodes	$\frac{1}{N(N-1)}\sum_{j\neq i}\delta_{ij}$
(Degree) Assortativity	Correlation between degrees of connected nodes	$\frac{\sum\limits_{(i,j)\in E}(D_i-\mu_\rho)(D_j-\mu_\rho)}{\sqrt{\sum\limits_{(i,j)\in E}(D_i-\mu_\rho)^2}\sqrt{\sum\limits_{(i,j)\in E}(D_j-\mu_\rho)^2}}$

1. We examine global clustering in the formula above.
2. For Average Path Length, δ_{ij} is the minimum distance between nodes i and j as defined in Table EC.1.
3. There are measures of assortativity based on homophily, e.g. gender or income assortativity. Here, we focus on degree assortativity.

$$l(\beta|X) \quad = \sum_{i=1}^{N} l_i(\beta|X_i) \tag{EC.2}$$

The adoption process is estimated by maximum likelihood estimation.

WOM Process

Given adoption parameters β, 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.

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 sampling algorithm. 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 sampling method (Random, Local or Global 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. The number of time periods varies across villages in the data, with a mean of 6.5 and SD of 1.83.

(1) At the beginning, a household that has become informed decides whether to adopt.

(2) Then, according to its adoption status, a 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 node has, as well as the adoption state $s \in \{A, NA\}$ of the informed node. We separate out the probabilities $p^{NA}(D)$ and $p^A(D)$ as detailed in §3.

(3) When this communication takes place, the 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 Informed (I) state.

For each simulation and for each village, we compute 7 moments according to Table EC.3 at the end of T periods of simulation. Thus, we have $N_{moments} = 7 \times 43 = 301$ moments across the villages with microfinance adoption. 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 7 in Section 4. For the MSM objective, we start with the initial weight matrix set to the identity matrix to obtain consistent estimates.

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 $q = (q_0^{NA}, q_1^{NA}, q_0^A, q_1^A) \in [0,1]^4$.

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 q.

(c): We use the cross-section and time series adoption status data to compute the moments detailed in Table EC.3 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.

Table EC.3 List of Moments.

Symbol	Description	Definition
MT1	Cumulative adoption upto time t (Time series moment)	$y_t = \frac{1}{N}\sum_{j=1}^{N} y_{jt}$
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	$cov\left(y, z(2)\right)$

(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.

EC.3. Moment Conditions for Estimation

In this section, we describe the rationales for the moments listed in Table EC.3 that we use in our estimation.

In general, all moments are informative in the estimation of all parameters. However, the connections between some moments and parameters are more intuitive. Moment MT1 is especially important for identification when there are differential effects for leaders. We describe the moments and the obvious associated links with parameters below.

(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 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.

(MT1) matches the cumulative overall adoption in each time period within each village. This is the typical data used in estimation of aggregate Bass-like diffusion models. The (MT1) moment 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.

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.

EC.4. Model Fit

Additional Model Fit Metrics

Next, we evaluate the fit of these models below using 3 additional measures. The metrics used for fit are as follows:

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 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 estimate 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.4 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 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
(SE)	0.02	0.02	0.02	0.02
Simulated Adoption	0.874	0.89	0.875	0.87
(SE)	0.097	0.098	0.096	0.1
RMSE	0.067	0.067	0.069	0.069
MAPE ($\times$100%)	0.379	0.372	0.395	0.406

Table EC.5 provides results for all the model specifications. We find that across the specifications, the models seem to be fairly similar in terms of their fit.

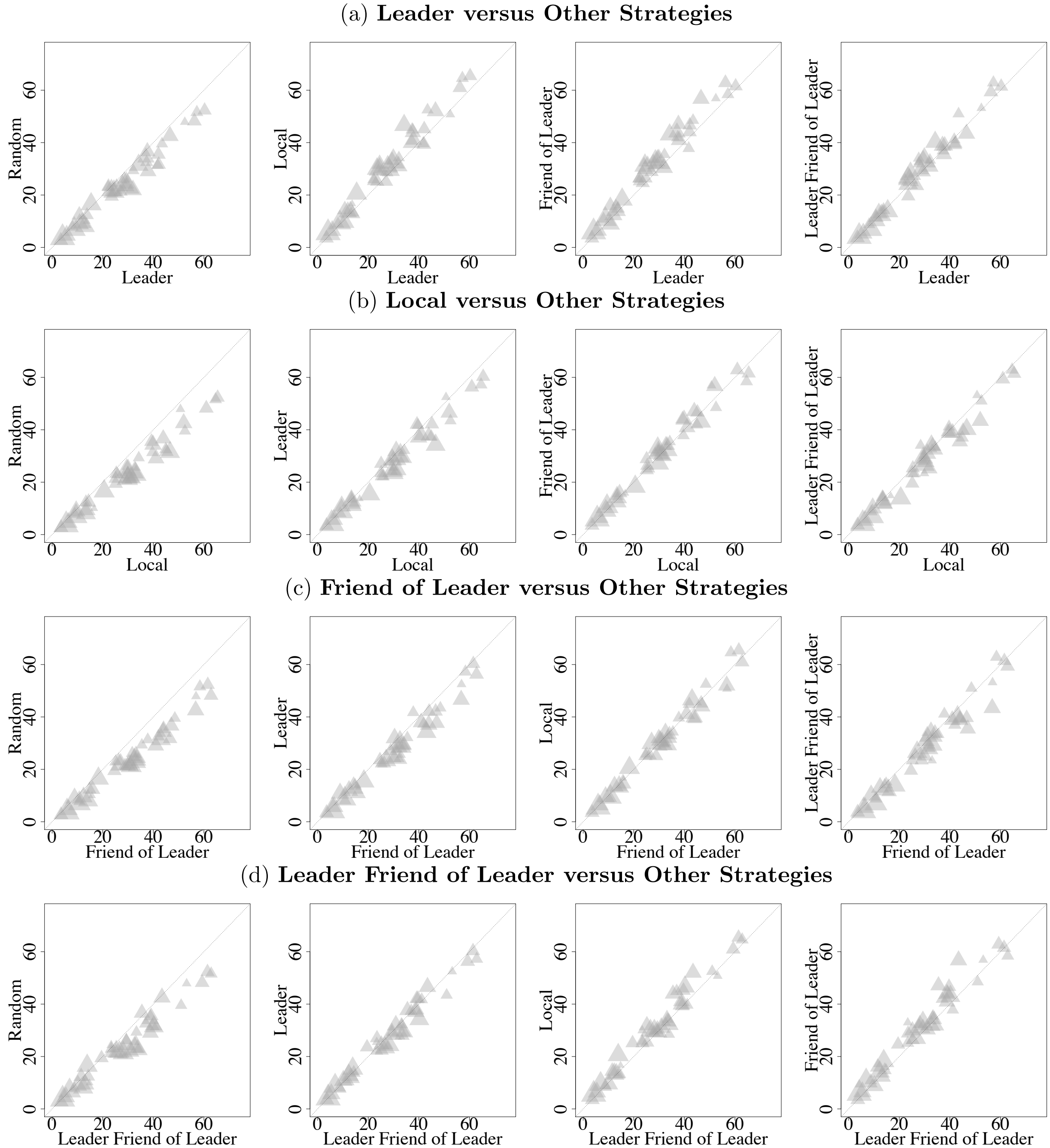
Table EC.5 Additional Fit Measures for Models

	$\mathbf{M_1^{E=0}}$	$\mathbf{M_2^{E=0}}$	$\mathbf{M_3^{E=0}}$	$\mathbf{M_4^{E=0}}$	$\mathbf{M_1^{E=1}}$	$\mathbf{M_2^{E=1}}$	$\mathbf{M_3^{E=1}}$	$\mathbf{M_4^{E=1}}$
Intercept	0.002	0.002	-0.003	-0.002	0.003	0.000	-0.004	0.001
(SE)	(0.021)	(0.02)	(0.02)	(0.021)	(0.019)	(0.02)	(0.02)	(0.021)
S.Adoption	0.877	0.874	0.901	0.899	0.867	0.89	0.904	0.876
(SE)	(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 ($\times$100%)	0.385	0.379	0.375	0.384	0.379	0.372	0.377	0.393

EC.5. Comparison of Strategies

In Figure EC.1, we plot the performance of the strategies pairwise, where performance is measured by the proportion of informed households in each counterfactual strategy evaluation.

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



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.

EC.6. Strategy Implementation

We detail the implementation of each of the strategies here in Table EC.6 below.

Table EC.6 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)
Friendship	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	Leader	Select node from list of leaders	List of Leaders (where leadership is specific to domain)
	Like Leader	Select leader node ℓ at random. Select the non-leader node most similar to ℓ in terms of network properties‡.	List of leaders + Entire Social Network
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
Other Strategies Not Examined in this Paper			
Influence	Most Central	Select most central node as seed	Complete List of Nodes + Degrees of *all* nodes + **Relevant Network** of Connections (Depends on domain of interest)[6]
	Most Influential	Compute Influence Score (e.g. Clout) for each node	List of Nodes + **Relevant Network** of Connections + Outcome variable to measure past influence + Attribution Mechanism

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

EC.7. 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 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 Tables EC.7 to EC.9.

A few observations are relevant here:

(a) Leader strategy always outperforms the random node strategy for any combination of model / (#seeds)

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

(c) The Hybrid seeding strategy achieves better performance than Local strategy in most model specifications. However, the leader hybrid strategy seems to consistently underperform the Leader strategy

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

(e) Leverage for all counterfactual strategies decreases as the number of seeds increases. Of course, in the limit where all nodes are chosen to be seeds, then all strategies perform equally well.

Table EC.7 Leverage for Counterfactual Strategies

Seeding at 5% of number of nodes

	$\mathbf{M_1^{E=0}}$	$\mathbf{M_2^{E=0}}$	$\mathbf{M_3^{E=0}}$	$\mathbf{M_4^{E=0}}$	$\mathbf{M_1^{E=1}}$	$\mathbf{M_2^{E=1}}$	$\mathbf{M_3^{E=1}}$	$\mathbf{M_4^{E=1}}$	$\mathbf{M_B^{E=0}}$	$\mathbf{M_B^{E=1}}$
Leader	1.069	1.064	1.070	1.064	1.069	1.068	1.059	1.070	1.050	1.066
Local Friend	1.128	1.116	1.120	1.118	1.116	1.122	1.120	1.123	1.091	1.108
Hybrid	1.121	1.119	1.119	1.118	1.126	1.126	1.121	1.127	1.086	1.114
Leader Hybrid	1.077	1.070	1.069	1.065	1.078	1.074	1.066	1.079	1.051	1.067
Like Leader	1.067	1.064	1.066	1.066	1.069	1.073	1.057	1.068	1.053	1.063

Table EC.8 Leverage for Counterfactual Strategies

Seeding at 1% of number of nodes

	$\mathbf{M_1^{E=0}}$	$\mathbf{M_2^{E=0}}$	$\mathbf{M_3^{E=0}}$	$\mathbf{M_4^{E=0}}$	$\mathbf{M_1^{E=1}}$	$\mathbf{M_2^{E=1}}$	$\mathbf{M_3^{E=1}}$	$\mathbf{M_4^{E=1}}$	$\mathbf{M_B^{E=0}}$	$\mathbf{M_B^{E=1}}$
Leader	1.121	1.119	1.125	1.136	1.103	1.112	1.133	1.116	1.108	1.111
Local Friend	1.202	1.210	1.198	1.210	1.224	1.236	1.222	1.206	1.178	1.195
Hybrid	1.230	1.235	1.207	1.226	1.218	1.241	1.238	1.209	1.203	1.203
Leader Hybrid	1.131	1.148	1.126	1.148	1.137	1.160	1.139	1.128	1.115	1.130
Like Leader	1.122	1.136	1.113	1.121	1.112	1.130	1.108	1.105	1.110	1.107

Table EC.9 Leverage for Counterfactual Strategies

Seeding at 0.5% of number of nodes

	$\mathbf{M_1^{E=0}}$	$\mathbf{M_2^{E=0}}$	$\mathbf{M_3^{E=0}}$	$\mathbf{M_4^{E=0}}$	$\mathbf{M_1^{E=1}}$	$\mathbf{M_2^{E=1}}$	$\mathbf{M_3^{E=1}}$	$\mathbf{M_4^{E=1}}$	$\mathbf{M_B^{E=0}}$	$\mathbf{M_B^{E=1}}$
Leader	1.164	1.160	1.135	1.161	1.137	1.123	1.145	1.155	1.142	1.153
Local Friend	1.299	1.293	1.272	1.283	1.282	1.247	1.283	1.256	1.255	1.266
Hybrid	1.310	1.309	1.261	1.292	1.318	1.294	1.310	1.302	1.256	1.281
Leader Hybrid	1.232	1.210	1.160	1.191	1.203	1.152	1.163	1.196	1.157	1.171
Like Leader	1.163	1.181	1.117	1.142	1.164	1.120	1.164	1.133	1.139	1.136