For Online Publication

Appendix B. Proofs

B.1. Relation of Diffusion Centrality to Other Measures.

We prove all of the statements for the case of weighted ($w_{ij} \in [0,1]$) and directed networks. Thus, $\mathbf{w} \in [0,1]^{n\times n}$ to allow for full heterogeneity in communication. For instance, w_{ij} and w_{ik} could both be positive, and yet differ from each other, or one could be positive and the other zero, or both zero, etc.

Let $v^{(L,k)}$ indicate k-th left-hand side eigenvector of $\mathbf{g}$ and similarly let $v^{(R,k)}$ indicate $\mathbf{g}$'s k-th right-hand side eigenvector. In the case of undirected networks, $v^{(L,k)} = v^{(R,k)}$.

Let $d(\mathbf{w})$ denote (out) degree centrality (so $d_i(\mathbf{w}) = \sum_j w_{ij}$). Eigenvector centrality corresponds to $v^{(R,1)}(\mathbf{w})$: the first eigenvector of $\mathbf{w}$. Also, let $GKB(\mathbf{w})$ denote generalized Katz–Bonacich centrality – defined for $\lambda_1(\mathbf{w}) < 1$ by:[36]

$$KB(\mathbf{w}) := \left(\sum_{t=1}^{\infty} (\mathbf{w})^t \right) \cdot \mathbf{1}.$$

It is direct to see that (i) diffusion centrality is proportional to out degree centrality at the extreme at which $T = 1$, and (ii) if $\lambda_1 < 1$, then diffusion centrality coincides with generalized Katz–Bonacich centrality if we set $T = \infty$. We now show that when $\lambda_1 > 1$ diffusion centrality approaches eigenvector centrality as T approaches ∞, which then completes the picture of the relationship between diffusion centrality and extreme centrality measures.

The difference between the extremes of Katz–Bonacich centrality and eigenvector centrality depends on whether λ_1 is sufficiently small so that limited diffusion takes place even in the limit of large T, or whether λ_1 is sufficiently large so that the knowledge saturates the network and then it is only relative amounts of saturation that are being measured.[37]

Theorem **B.1.**

[36]See (2.7) in Jackson (2008b) for additional discussion and background. This is a generalization of a measure first discussed by Katz, and corresponds to Bonacich's definition if the network is unweighted and all passing probabilities are the same, and then both of Bonacich's parameters are set to q.

[37]Saturation occurs when the entries of $\left(\sum_{t=1}^{\infty} (\mathbf{w})^t \right) \cdot \mathbf{1}$ diverge (note that in a [strongly] connected network, if one entry diverges, then all entries diverge). Nonetheless, the limit vector is still proportional to a well defined limit vector: the first eigenvector.

(1) *Diffusion centrality is proportional to (out) degree when $T = 1$:*

$$DC(\mathbf{w}, 1) = d(\mathbf{w}).$$

(2) *If $\lambda_1 \geq 1$ and $\mathbf{w}$ is aperiodic, then as $T \to \infty$ diffusion centrality approximates eigenvector centrality:*

$$\lim_{T\to\infty} \frac{DC(\mathbf{w}, T)}{\sum_{t=1}^{T} (\lambda_1)^t} = v^{(R,1)}.$$

(3) *For $T = \infty$ and $\lambda_1 < 1$, diffusion centrality is Generalized Katz–Bonacich centrality:*

$$DC(\mathbf{w}, \infty) = KB(\mathbf{w}).$$

This is a result we mention in Banerjee, Chandrasekhar, Duflo, and Jackson (2013). An independent formalization appears in Benzi and Klymko (2014).

We also remark on the comparison to another measure: the influence vector that appears in the DeGroot learning model (see, e.g., Golub and Jackson (2010)). That metric captures how influential a node is in a process of social learning. It corresponds to the (left-hand) unit eigenvector of a stochasticized matrix of interactions rather than a raw adjacency matrix. While it might be tempting to use that metric here as well, we note that it is the right conceptual object to use in a process of *repeated averaging* through which individuals update opinions based on averages of their neighbors' opinions. It is thus conceptually different from the diffusion process that we study. Nonetheless, one can also define a variant of diffusion centrality that works for finite iterations of DeGroot updating.

Proof of Theorem B.1. We show the second statement as the others follow directly.

First, consider any irreducible and aperiodic nonnegative (and hence primitive) $\mathbf{w}$. If the statement holds for any arbitrarily close positive and diagonalizable $\mathbf{w}'$ (which are dense in a nonnegative neighborhood of $\mathbf{w}$), then since $\frac{DC(\mathbf{w},T)}{\sum_{t=1}^{T}(\lambda_1)^t}$ is a continuous function (in a neighborhood of a primitive $\mathbf{w}$, which has a simple first eigenvalue), as is the first eigenvector, then the statement also holds at $\mathbf{w}$.[38] Thus, it is enough to prove the result for a positive and diagonalizable $\mathbf{w}$.

We show the following for a positive and diagonalizable $\mathbf{w}$:

[38]As is shown below, $\frac{DC(\mathbf{w},T)}{\sum_{t=1}^{T}(\lambda_1)^t}$ has a well-defined limit, and so this holds also for the limit.

- If $\lambda_1 > 1$, then
$$\lim_{T\to\infty} \frac{DC(\mathbf{w},T)}{\sum_{t=1}^{T} (\lambda_1)^t} = \lim_{T\to\infty} \frac{DC(\mathbf{w},T)}{\frac{\lambda_1-(\lambda_1)^{T+1}}{1-(\lambda_1)}} = v^{(R,1)}.$$
- If $\lambda_1 = 1$, then
$$\lim_{T\to\infty} \frac{1}{T} DC(\mathbf{w},T) = v^{(R,1)}.$$

Normalize the eigenvectors to lie in ℓ_1, so that the entries in each column of $\mathbf{V}^{-1}$ and each row of $\mathbf{V}$ sum to 1.

Let us first show the statement for the case where $\lambda_1 = 1$. It is sufficient to show
$$\lim_{T\to\infty} \left\| \frac{DC(\mathbf{w},T)}{T} - v^{(R,1)} \right\| = 0.$$
First, note that given the diagonalizable matrix, straightforward calculations show that
$$DC_i(\mathbf{w},T) = \sum_j \sum_{t=1}^{T} \sum_k v_i^{(R,k)} v_j^{(L,k)} \lambda_k^t.$$
Thus,
$$\begin{aligned}
\left| \frac{DC_i(\mathbf{w},T)}{T} - v_i^{(R,1)} \right| &= \left| \frac{\sum_j \sum_{t=1}^{T} \sum_{k=1}^{n} v_i^{(R,k)} v_j^{(L,k)} \lambda_k^t}{T} - v_i^{(R,1)} \right| = \\
&= \left| \frac{1}{T} \sum_j \sum_{t=1}^{T} \sum_{k=2}^{n} v_i^{(R,k)} v_j^{(L,k)} \lambda_k^t \right| \leq \frac{1}{T} \sum_{t=1}^{T} \sum_{k=2}^{n} 1 \cdot \underbrace{\left| \sum_{j=1}^{n} v_j^{(L,k)} \right|}_{\leq 1} \cdot \left| \lambda_k^t \right| \\
&\leq \frac{n}{T} \sum_{t=1}^{T} \left| \lambda_2^t \right| = \frac{n}{T} \frac{|\lambda_2|}{1-|\lambda_2|} \left(1 - |\lambda_2|^T\right) \to 0.
\end{aligned}$$
Since the length of the vector (which is n) is unchanging in T, pointwise convergence implies convergence in norm, proving the result.

The final piece repeats the argument for $\lambda_1 > 1$. We show
$$\lim_{T\to\infty} \left\| \frac{DC(\mathbf{w},T)}{\sum_{t=1}^{T} (\lambda_1)^t} - v^{(R,1)} \right\| = 0.$$

By similar derivations as above,

$$\begin{aligned}
\left|\frac{DC_i\left(\mathbf{w},T\right)}{\sum_{t=1}^T \lambda_1^t} - v_i^{(R,1)}\right| &= \left|\frac{\sum_j \sum_{t=1}^T \sum_{k=1}^n v_i^{(R,k)} v_j^{(L,k)} \lambda_k^t}{\sum_{t=1}^T \lambda_1^t} - v_i^{(R,1)}\right| \\
&= \left|\frac{\sum_j \sum_{t=1}^T \sum_{k=2}^n v_i^{(R,k)} v_j^{(L,k)} \lambda_k^t}{\sum_{t=1}^T \lambda_1^t} + \frac{\sum_j \sum_{t=1}^T v_i^{(R,1)} v_j^{(L,1)} \lambda_1^t}{\sum_{t=1}^T \lambda_1^t} - v_i^{(R,1)}\right| \\
&= \left|\frac{\sum_j \sum_{t=1}^T \sum_{k=2}^n v_i^{(R,k)} v_j^{(L,k)} \lambda_k^t}{\sum_{t=1}^T \lambda_1^t} + \frac{\sum_{t=1}^T v_i^{(R,1)} \lambda_1^t}{\sum_{t=1}^T \lambda_1^t} - v_i^{(R,1)}\right| \\
&= \left|\frac{1}{\sum_{t=1}^T \lambda_1^t} \sum_j \sum_{t=1}^T \sum_{k=2}^n v_i^{(R,k)} v_j^{(L,k)} \lambda_k^t\right| \\
&\leq \frac{1}{\sum_{t=1}^T \lambda_1^t} \sum_{t=1}^T \sum_{k=2}^n 1 \cdot \left|\sum_{j=1}^n v_j^{(L,k)}\right| \cdot \left|\lambda_k^t\right| \\
&\leq \frac{n}{\sum_{t=1}^T \lambda_1^t} \sum_{t=1}^T \left|\lambda_2^t\right|.
\end{aligned}$$

Note that this last expression converges to 0 since $\lambda_1 > 1$, and $\lambda_1 > \lambda_2$.[39] which completes the argument. ∎

B.2. **Other Proofs.**

Proof of Theorem A.1 .

$$\begin{aligned}
\mathrm{E}\left[DC\left(\mathbf{g}(n,p);q,T\right)\right]_i &= \left[\sum_1^T \mathrm{E}\left[q^t \mathbf{g}(n,p)^t\right] \cdot \mathbf{1}\right]_i \\
&= \sum_1^T q^t n \mathrm{E}\left[\mathbf{g}(n,p)^t\right]_{ij},
\end{aligned}$$

where the last equality comes from the fact that $\mathrm{E}\left[\mathbf{g}(n,p)^t\right]_{ij} = \mathrm{E}\left[\mathbf{g}(n,p)^t\right]_{ik}$ for all i,j,k in an Erdos–Renyi random graph.

Next, note that

$$\mathrm{E}\left[\mathbf{g}(n,p)^t\right]_{ij} = \mathrm{E}\left[\sum_{k_1,k_2,\ldots,k_{t-1}\in\{1,\ldots,n\}^{t-1}} g_{ik_1} g_{k_1k_2} \cdots g_{k_{t-1}j}\right]$$

[39] Note that it is important that $\lambda_1 \geq 1$ for this claim. In that case, observe that

$$\frac{\sum_{t=1}^T |\lambda_2|^t}{\sum_{t=1}^T \lambda_1^t} = \frac{\lambda_2}{\lambda_1} \cdot \frac{1-\lambda_1}{1-\lambda_2}$$

by the properties of a geometric sum, which is of constant order. Thus, higher order terms (λ_2, etc.) persistently matter and are not dominated relative to $\sum_t^T \lambda_1^t$.

If all the indexed $g_{..}$'s were distinct, the right hand side of this equation would simply be $n^{t-1}p^t$. However, in the summand sometimes terms repeat. For example, if there were exactly x repetitions, the probability of getting the walk would be p^{t-x} instead of p^t. Thus, it follows directly that

$$\mathrm{E}\left[\mathbf{g}(n,p)^t\right]_{ij} \geq n^{t-1}p^t$$

and so

$$\begin{aligned}\mathrm{E}\left[DC\left(\mathbf{g}(n,p);q,T\right)\right]_i &= \sum_1^T q^t n \mathrm{E}\left[\mathbf{g}(n,p)^t\right]_{ij} \\ &\geq \sum_1^T q^t n^t p^t = npq\frac{1-(npq)^T}{1-npq}\end{aligned}$$

Note also, that

$$\mathrm{E}\left[\sum_{k_1,k_2,\ldots,k_t\in\{1,\ldots,n\}^t} g_{ik_1}g_{k_1k_2}\cdots g_{k_{t-1}j}\right] \leq n^{t-1}p^t + tn^{t-2}p^{t-1} + t^2n^{t-3}p^{t-2} + \ldots + t^t.$$

This last inequality is a very loose upper bound generated by setting a loose upper bound on how many $g_{..}$'s could conceivably repeat, and then putting in the expression that would ensue if they did repeat. Despite how loose the bound is, it suffices for our purposes.

Given that $t \leq T < pn$, it follows that

$$\begin{aligned}\mathrm{E}\left[\sum_{k_1,k_2,\ldots,k_t\in\{1,\ldots,n\}^t} g_{ik_1}g_{k_1k_2}\cdots g_{k_{t-1}j}\right] &\leq n^{t-1}p^t\left(1+\frac{t}{pn}+\left(\frac{t}{pn}\right)^2\ldots+\left(\frac{t}{pn}\right)^t\right) \\ &= n^{t-1}p^t\left(\frac{1-\left(\frac{t}{pn}\right)^t}{1-\left(\frac{t}{pn}\right)}\right).\end{aligned}$$

Thus,

$$\mathrm{E}\left[\mathbf{g}(n,p)^t\right]_{ij} \leq n^{t-1}p^t\frac{1}{1-\frac{T}{pn}}.$$

Since $T << pn$ it follows that (here $o(1)$ is with respect to n):

$$\begin{aligned}\mathrm{E}\left[DC\left(\mathbf{g}(n,p);q,T\right)\right]_i &= \sum_1^T q^t n \mathrm{E}\left[\mathbf{g}(n,p)^t\right]_{ij} \\ &\leq \sum_1^T q^t n^t p^t(1+o(1)) = npq\frac{1-(npq)^T}{1-npq}(1+o(1)).\end{aligned}$$

The theorem follows directly. ■

Proof of Theorem 1 . Recall that $\mathbf{H} = \sum_{t=1}^{T} (\mathbf{w})^t$ and $DC = \left(\sum_{t=1}^{T} (\mathbf{w})^t\right) \cdot \mathbf{1}$ and so

$$DC_i = \sum_j H_{ij}.$$

Additionally,

$$\mathrm{cov}(DC, H_{\cdot,j}) = \sum_i \left(DC_i - \sum_k \frac{DC_k}{n}\right)\left(H_{ij} - \sum_k \frac{H_{kj}}{n}\right).$$

Thus

$$\sum_j \mathrm{cov}(DC, H_{\cdot,j}) = \sum_i \left(DC_i - \sum_k \frac{DC_k}{n}\right)\left(\sum_j H_{ij} - \sum_k \frac{\sum_j H_{kj}}{n}\right),$$

implying

$$\sum_j \mathrm{cov}(DC, H_{\cdot,j}) = \sum_i \left(DC_i - \sum_k \frac{DC_k}{n}\right)\left(DC_i - \sum_k \frac{DC_k}{n}\right) = \mathrm{var}(DC),$$

which completes the proof. ■

Proof of Corollary A.1 . To see (1), first note that $x\frac{1-x^T}{1-x} \to 0$ if $x \to 0$, and that $x\frac{1-x^T}{1-x} \to x\frac{x^T}{x} \to \infty$ if $x \to \infty$. Replacing x with npq and then applying Theorem A.1 yields the result under (a) and (b), respectively.

To see (2), we consider the case in which $q > 1/(\mathrm{E}[\lambda_1])^{1-\varepsilon}$, which of course is equivalent to $npq > (np)^{\varepsilon}$. This is the case under which (b) applies. This also implies the result in (a), since if the conclusion of (a) holds for such a q it will also hold for all lower q, given that DC is monotone in q.

Again, since $npq > 1$, it follows that if T is growing, then

$$\mathrm{E}\left[DC\left(q\mathbf{g}(n,p),T\right)\right]_i \to npq\frac{1-(npq)^T}{1-npq} \to (npq)^T.$$

So, to have

$$\mathrm{E}\left[DC\left(q\mathbf{g}(n,p),T\right)\right]_i \geq kn$$

for some $k > 0$, it is sufficient that $(npq)^T \geq kn$, or

$$T \geq \frac{\log(n)+\log(k)}{\log np + \log(q)} \to \frac{\log(n)}{\log np} \sim \mathrm{E}[Diam\left(\mathbf{g}(n,p)\right)],$$

where the last comparison is a property of Erdos–Renyi random networks given that $\frac{1-\varepsilon}{\sqrt{n}} \geq p \geq (1+\varepsilon)\frac{\log(n)}{n}$, and so this establishes (b). From the analogous calculation, if T is below $\frac{\log(n)}{\log np}$, then $\mathrm{E}\left[DC\left(q\mathbf{g}(n,p),T\right)\right]_i \leq kn$ for any k, and so (a) follows. ■

Proof of Theorem 2. Again, we prove the result for a positive diagonalizable $\mathbf{w}$, noting that it then holds for any (nonnegative) $\mathbf{w}$.

Again, let $\mathbf{w}$ be written as

$$\mathbf{w} = \mathbf{V}\Lambda\mathbf{V}^{-1}.$$

It then follows that we can write

$$\mathbf{H} = \sum_{t=1}^{T} (\mathbf{w})^t = \sum_{t=1}^{T} \left(\sum_{k=1}^{n} v_i^{(R,k)} v_j^{(L,k)} \lambda_k^t \right).$$

By the ordering of the eigenvalues from largest to smallest in magnitude,

$$\begin{aligned} \mathbf{H}_{\cdot,j} &= \sum_{t=1}^{T} \left[v^{(R,1)} v_j^{(L,1)} \lambda_1^t + v^{(R,2)} v_j^{(L,2)} \lambda_2^t + O\left(|\lambda_2|^t \right) \right] \\ &= \sum_{t=1}^{T} \left[v^{(R,1)} v_j^{(L,1)} \lambda_1^t + O\left(|\lambda_2|^t \right) \right] \\ &= v^{(R,1)} v_j^{(L,1)} \sum_{t=1}^{T} \lambda_1^t + O\left(\sum_{t=1}^{T} |\lambda_2|^t \right). \end{aligned}$$

So, since the largest eigenvalue is unique, it follows that

$$\frac{\mathbf{H}_{\cdot,j}}{\sum_{t=1}^{T} \lambda_1^t} = v^{(R,1)} v_j^{(L,1)} + O\left(\frac{\sum_{t=1}^{T} |\lambda_2|^t}{\sum_{t=1}^{T} \lambda_1^t} \right).$$

Note that the last expression converges to 0 since $\lambda_1 > 1$, and $\lambda_1 > \lambda_2$. Thus,

$$\frac{\mathbf{H}_{\cdot,j}}{\sum_{t=1}^{T} \lambda_1^t} \to v^{(R,1)} v_j^{(L,1)}$$

for each j. This completes the proof since each column of $\mathbf{H}$ is proportional to $v^{(R,1)}$ in the limit, and thus has the correct ranking for large enough T.[40] Note that the ranking is up to ties, as the ranking of tied entries may vary arbitrarily along the sequence. That is, if $v_i^{(R,1)} = v_\ell^{(R,1)}$, then j's ranking over i and ℓ could vary arbitrarily with T, but their rankings will be correct relative to any other entries with higher or lower eigenvector centralities. ∎

[40] The discussion in Footnote 39 clarifies why $\lambda_1 > 1$ is required for the argument.

Appendix C. Extension of Microfinance Village (wave 2) Network results

This section extends the descriptive analysis from the Microfinance Village (wave 2) network data on 33 villages. We repeat all of our analyses with OLS specifications instead of Poisson specifications. Additionally, we include a Post-LASSO estimation which conducts a LASSO to select which variables best explain our outcome of interest (number of nominations) and then does a post-estimation to recover consistent parameter estimates.

Table C.1. Factors predicting nominations

	(1) Event	(2) Event	(3) Event	(4) Event	(5) Event
Diffusion Centrality	0.285				
	(0.060)				
Degree Centrality		0.250			
		(0.061)			
Eigenvector Centrality			0.283		
			(0.064)		
Leader				0.436	
				(0.168)	
Geographic Centrality					-0.025
					(0.038)
Observations	6,466	6,466	6,466	6,466	6,466
	(1) Loan	(2) Loan	(3) Loan	(4) Loan	(5) Loan
Diffusion Centrality	0.391				
	(0.071)				
Degree Centrality		0.367			
		(0.065)			
Eigenvector Centrality			0.378		
			(0.074)		
Leader				0.653	
				(0.224)	
Geographic Centrality					-0.045
					(0.029)
Observations	6,466	6,466	6,466	6,466	6,466

Notes: This table uses data from the microfinance village (wave 2) dataset. It reports estimates of OLS regressions where the outcome variable is the expected number of nominations under the event question. Panel A presents results for the event question, and Panel B presents results for the loan question. Degree centrality, eigenvector centrality, and diffusion centrality, $DC\left(1/\mathrm{E}[\lambda_1]\mathbf{g}, \mathrm{E}[Diam(\mathbf{g}(n,p))]\right)$, are normalized by their standard deviations. Standard errors (clustered at the village level) are reported in parentheses.

TABLE C.2. Factors predicting nominations

	(1) Event	(2) Event	(3) Event	(4) Event	(5) Event	(6) Event
Diffusion Centrality	0.303	0.161	0.269	0.285	0.173	0.285
	(0.091)	(0.087)	(0.061)	(0.060)	(0.107)	(0.060)
Degree Centrality	-0.020				-0.013	
	(0.066)				(0.068)	
Eigenvector Centrality		0.138			0.137	
		(0.095)			(0.095)	
Leader			0.294			
			(0.174)			
Geographic Centrality				-0.026		
				(0.039)		
Observations	6,466	6,466	6,466	6,466	6,466	6,466
Post-LASSO						✓

	(1) Loan	(2) Loan	(3) Loan	(4) Loan	(5) Loan	(6) Loan
Diffusion Centrality	0.310	0.266	0.366	0.391	0.175	0.310
	(0.112)	(0.089)	(0.071)	(0.071)	(0.124)	(0.112)
Degree Centrality	0.091				0.098	0.091
	(0.079)				(0.079)	(0.079)
Eigenvector Centrality		0.138			0.144	
		(0.089)			(0.087)	
Leader			0.461			
			(0.229)			
Geographic Centrality				-0.045		
				(0.030)		
Observations	6,466	6,466	6,466	6,466	6,466	6,466
Post-LASSO						✓

Notes: This table uses data from the microfinance village (wave 2) dataset. It reports estimates of OLS regressions where the outcome variable is the expected number of nominations. Panel A presents results for the event question, and Panel B presents results for the loan question. Degree centrality, eigenvector centrality, and diffusion centrality, $DC\,(1/\mathrm{E}[\lambda_1]\mathbf{g}, \mathrm{E}[Diam(\mathbf{g}(n,p))])$, are normalized by their standard deviations. Column (6) uses a post-LASSO procedure where in the first stage LASSO is implemented to select regressors and in the second stage the regression in question is run on those regressors. Omitted terms indicate they were not selected in the first stage. Standard errors (clustered at the village level) are reported in parentheses.

Appendix D. Extension of experiment analysis

This section extends the analysis of the experiment results to using four instruments.

Table D.1. Calls received by treatment

	(1) RF Calls Received	(2) OLS Calls Received	(3) IV 1: First Stage At least 1 Gossip	(4) IV 2: First Stage At least 1 Elder	(5) IV: Second Stage Calls Received
Gossip Treatment	4.559		0.795	0.430	
	(3.121)		(0.0753)	(0.108)	
5 Gossip Seeds	-1.785		-0.303	-0.206	
	(5.290)		(0.110)	(0.153)	
Elder Treatment	2.279		0.370	0.872	
	(2.424)		(0.106)	(0.0685)	
5 Elder Seeds	-6.798		-0.272	-0.0578	
	(3.487)		(0.149)	(0.100)	
At least 1 Gossip		3.786			8.063
		(1.858)			(3.845)
At least 1 Elder		0.792			-3.684
		(2.056)			(2.266)
Observations	212	212	212	212	212
Control Group Mean	8.019	5.846	0.389	0.183	5.496

	(1) RF $\frac{\text{Calls Received}}{\text{Seeds}}$	(2) OLS $\frac{\text{Calls Received}}{\text{Seeds}}$	(3) IV 1: First Stage At least 1 Gossip	(4) IV 2: First Stage At least 1 Elder	(5) IV: Second Stage $\frac{\text{Calls Received}}{\text{Seeds}}$
Gossip Treatment	1.593		0.795	0.430	
	(1.030)		(0.0753)	(0.108)	
5 Gossip Seeds	-1.083		-0.303	-0.206	
	(1.348)		(0.110)	(0.153)	
Elder Treatment	0.622		0.370	0.872	
	(0.770)		(0.106)	(0.0685)	
5 Elder Seeds	-1.430		-0.272	-0.0578	
	(0.912)		(0.149)	(0.100)	
At least 1 Gossip		0.952			2.169
		(0.501)			(1.043)
At least 1 Elder		0.309			-0.676
		(0.511)			(0.578)
Observations	212	212	212	212	212
Control Group Mean	1.953	1.451	0.389	0.183	1.186

Notes: This table uses data from the cell phone RCT dataset. Panel A uses the number of calls received as the outcome variable. Panel B normalizes the number of calls received by the number of seeds, 3 or 5, which is randomly assigned. For both panels, Column (1) shows the reduced form results of regressing number of calls received on dummies for gossip treatment and elder treatment. Column (2) regresses number of calls received on the dummies for if at least 1 gossip was hit and for if at least 1 elder was hit in the village. Columns (3) and (4) show the first stages of the instrumental variable regressions, where the dummies for "at least 1 gossip" and "at least 1 elder" are regressed on the exogenous variables: gossip treatment dummy, 5 gossip seeds dummy, elder treatment dummy, 5 elder seeds dummy. Column (5) shows the second stage of the IV; it regresses the number of calls received on the dummies for if at least 1 gossip was hit and if at least 1 elder was hit, both instrumented by treatment status of the village (gossip treatment or not, elder treatment or not) and seed number dummies for the village (5 gossip seeds or not, 5 elder seeds or not). All columns control for number of gossips, number of elders and number of seeds. For columns (1), (3), and (4) the control group mean is calculated as the mean expectation of the outcome variable when the treatment is "random". For columns (2) and (5), the control group mean is calculated as the mean expectation of the outcome variable when no gossips or elders are reached. The control group mean for the second stage IV is calculated using IV estimates. Robust standard errors are reported in parentheses.

Appendix E. Experiment Analysis with Broadcast Village

This section repeats our main experimental analyses but includes the broadcast village where the poster was made by one of the seeds.

Table E.1. Calls received by treatment

VARIABLES	(1) RF Calls Received	(2) OLS Calls Received	(3) IV 1: First Stage At least 1 Gossip	(4) IV 2: First Stage At least 1 Elder	(5) IV: Second Stage Calls Received
Gossip Treatment	2.266		0.636	0.331	
	(3.116)		(0.0660)	(0.0821)	
Elder Treatment	-2.809		0.220	0.846	
	(2.577)		(0.0807)	(0.0502)	
At least 1 Gossip		5.005			6.122
		(2.210)			(4.532)
At least 1 Elder		-0.619			-4.914
		(2.472)			(2.628)
Observations	213	213	213	213	213
Control Group Mean	9.534	6.277	0.400	0.180	7.971
Gossip Treatment=Elder Treatment (pval.)	0.0300		0	0	
At least 1 Gossip=At least 1 Elder (pval.)		0.160			0.0300

VARIABLES	(1) RF $\frac{\text{Calls Received}}{\text{Seeds}}$	(2) OLS $\frac{\text{Calls Received}}{\text{Seeds}}$	(3) IV 1: First Stage At least 1 Gossip	(4) IV 2: First Stage At least 1 Elder	(5) IV: Second Stage $\frac{\text{Calls Received}}{\text{Seeds}}$
Gossip Treatment	0.591		0.636	0.331	
	(0.841)		(0.0660)	(0.0821)	
Elder Treatment	-0.646		0.220	0.846	
	(0.738)		(0.0807)	(0.0502)	
At least 1 Gossip		1.359			1.535
		(0.644)			(1.179)
At least 1 Elder		-0.162			-1.164
		(0.691)			(0.748)
Constant				0.109	
				(0.160)	
Observations	213	213	213	213	213
Control Group Mean	2.452	1.595	0.400	0.180	2.048
Gossip Treatment=Elder Treatment (pval.)	0.0400		0	0	
At least 1 Gossip=At least 1 Elder (pval.)		0.190			0.0400

Notes: This table uses data from the cell phone RCT dataset. Panel A uses the number of calls received as the outcome variable. Panel B normalizes the number of calls received by the number of seeds, 3 or 5, which is randomly assigned. For both panels, Column (1) shows the reduced form results of regressing number of calls received on dummies for gossip treatment and elder treatment. Column (2) regresses number of calls received on the dummies for if at least 1 gossip was hit and for if at least 1 elder was hit in the village. Columns (3) and (4) show the first stages of the instrumental variable regressions, where the dummies for "at least 1 gossip" and "at least 1 elder" are regressed on the exogenous variables: gossip treatment dummy and elder treatment dummy. Column (5) shows the second stage of the IV; it regresses the number of calls received on the dummies for if at least 1 gossip was hit and if at least 1 elder was hit, both instrumented by treatment status of the village (gossip treatment or not, elder treatment or not). All columns control for number of gossips, number of elders, and number of seeds. For columns (1), (3), and (4) the control group mean is calculated as the mean expectation of the outcome variable when the treatment is "random". For columns (2) and (5), the control group mean is calculated as the mean expectation of the outcome variable when no gossips or elders are reached. The control group mean for the second stage IV is calculated using IV estimates. Robust standard errors are reported in parentheses.

Table E.2. Calls received by seed type

VARIABLES	(1) Calls Received	(2) Calls Received	(3) Calls Received	(4) $\frac{\text{Calls Received}}{\text{Seeds}}$	(5) $\frac{\text{Calls Received}}{\text{Seeds}}$	(6) $\frac{\text{Calls Received}}{\text{Seeds}}$
At least 1 Gossip	12.89	13.02		3.751	3.871	
	(7.225)	(8.157)		(2.282)	(2.584)	
At least 1 Elder	-3.371	-3.321		-1.012	-0.962	
	(5.155)	(4.946)		(1.547)	(1.456)	
At least 1 High DC Seed		-0.485	2.262		-0.478	0.342
		(4.803)	(3.834)		(1.515)	(1.189)
Observations	69	69	69	69	69	69
Control Group Mean	4.840	4.840	8.828	1.101	1.101	2.433
At least 1 Gossip=At least 1 Elder (pval.)	0.150	0.170		0.190	0.200	
At least 1 Gossip=At least 1 High DC Seed (pval.)		0.270			0.270	
At least 1 Elder=At least 1 High DC Seed (pval.)		0.580			0.710	

Notes: This table uses data from the cell phone RCT and follow-up network dataset. The table presents OLS regressions of number of calls received (and number of calls received normalized by the number of seeds, 3 or 5, which is randomly assigned) on characteristics of the set of seeds. High DC refers to a seed being above the mean by one standard deviation of the centrality distribution. All columns control for total number of gossips, number of elders, and number of seeds. For columns (1), (2), (4), and (5), the control group mean is calculated as the mean expectation of the outcome variable when no gossips or elders are reached. For columns (3) and (6), the control group mean is calculated as the mean expectation of the outcome variable when no high DC seeds are reached. Robust standard errors are reported in parentheses.

Appendix F. When People Don't Nominate Anyone

Here we look at whether the odds that someone refuses to nominate anyone can be explained by our model.

Suppose that people get disutility from reporting incorrect guesses to a surveyor.[41]

Our theoretical results on network gossip converging to diffusion centrality are asymptotic results - with any finite number of iterations of gossip passing, an individual's ranking of others' centralities can be noisy.

In particular, after T periods, i's perceptions are not quite at their expectations:

$$\widehat{NG}_{ji} = NG_{ji} + \varepsilon_{ji}.$$

where ε_{ji} will depend on T, $\mathbf{w}$, and i and j's positions in the network - eventually vanishing (in proportion to NG_{ji}) when T becomes large.

One way that i can assess how accurate their ranking is, would be looking at how extreme $\widehat{NG}_{ji}$ is compared to the average number of times that i has heard about other people. If this is more extreme, then it is less likely to reverse. Then if i' 99th percentile $\widehat{NG}_{ji}$ is more extreme in terms of numbers of times heard compared to the average, then i can distinguish this tail better from the rest of the distribution, even if there is noise. That is, the ability to distinguish $NG_{ji}^{99\%}$ from $\bar{NG}_i$ should increase in $NG_{ji}^{99\%}$ holding ε fixed.

From this perspective, our model predicts that the larger the extreme quantiles of i's network gossip distribution are, controlling for the average (note that is just proportional to DC_i), then i should be more likely to nominate and less likely to answer that he does not know. This comes from the fact that he should be better able to distinguish between alternatives.

For this estimation, as in the rest of the paper we work with T equal to the diameter of the network and q equal to the inverse of the first eigenvalue.

The results in Table F.1 are consistent with this story. A one standard deviation increase in the 99th percentile of network gossip of j as perceived by i corresponds to, holding fixed the average network gossip of others as perceived by i, a 1.3 percent increase in the probability of i nominating anyone. Across specifications we have p values of $0.11, 0.08, 0.18, 0.6$, respectively, across columns.

[41]See Alatas et al. (2014) for such an example of where this happened in practice when individuals were more likely to report that they don't know rather than offer a guess when trying to rank other villagers in terms of wealth.

TABLE F.1. Does the tail of network gossip drive nominations?

VARIABLES	(1) Nominated Anyone	(2) Nominated Anyone	(3) Nominated Anyone	(4) Nominated Anyone
99th percentile of NG_{ji}	3.739	8.134		
	(2.348)	(4.651)		
98th percentile of NG_{ji}			3.499	6.174
			(2.620)	(4.409)
99th Percentile	✓	✓		
Village FE		✓		✓
98th Percentile			✓	✓

Notes: This table uses data from the microfinance village (wave 2) dataset. The data consists of a individual level observations and the outcome is whether the individual nominated anyone in response to the lottery gossip question.The key regressor is the value of the person who is at the 99th (or 9th) percentile from the distribution network gossip for i. Columns 2 and 4 include village fixed effects, estimated by a conditional logit. Standard errors (clustered at the village level) are reported in parentheses.

Next we look at the demographics of those who choose to nominate versus those who choose not to nominate in Table F.2. We look at a number of demographics: caste, occupation, household amenities, respondent gender, and geography. We also include social status variables such as leadership (as defined from our previous work based on the lender's designations) and the number of nominations the individual's household received under loan and event questions. Finally we include the diffusion centrality of the respondent's household.

Table F.2 presents the result. Being more diffusion central is positively and significantly associated with the respondent being willing to nominate someone. This is consistent with our model. Meanwhile, almost none of the other demographics or social status variables matter. In fact, there is no other statistically significant variable for loan nominations. For event nominations, the number of times the respondent's household was nominated under the event question matters as does whether the respondent owns their own house. Beyond that, no other variable matters.

Table F.2. Demographics of those who choose to nominate

VARIABLES	(1) Nominates Someone (Loan)	(2) Nominates Someone (Event)
Diffusion Centrality (Standardized)	0.024	0.015
	(0.008)	(0.008)
No. of Nominations (Loans)	-0.000	-0.001
	(0.003)	(0.003)
No. of Nominations (Events)	0.004	0.008
	(0.003)	(0.003)
Leader	0.004	-0.007
	(0.020)	(0.018)
SCST	-0.010	0.006
	(0.026)	(0.022)
Electrified	-0.031	-0.002
	(0.031)	(0.028)
Private Electrification	0.013	-0.003
	(0.017)	(0.017)
Own House	-0.036	-0.049
	(0.026)	(0.028)
No. of Rooms	-0.002	0.002
	(0.005)	(0.005)
Land Owner	-0.020	-0.013
	(0.028)	(0.026)
Farm Laborer	-0.032	-0.014
	(0.022)	(0.021)
Business Owner	-0.020	-0.014
	(0.027)	(0.025)
GPS Centrality	-0.007	-0.008
	(0.008)	(0.006)
Female Respondent	0.009	0.014
	(0.015)	(0.013)
Observations	5,707	5,707

Notes: This table uses data from the microfinance village (wave 2) dataset. The data consists of a individual level observations and the outcome is whether the individual nominated anyone in response to the lottery gossip question. Standard errors (clustered at the village level) are reported in parentheses and all specifications include village fixed effects.

Appendix G. Characteristics of Gossips, Elders, and Random

Table G.1. Characteristics of gossip, elder, and random households

VARIABLES	(1) SCST	(2) Laborer	(3) Land Owner	(4) Electrified	(5) Private Electricity	(6) Own House	(7) No. of Rooms
Gossip Nominee	-0.0278	-0.0729	0.0793	0.0173	0.0455	0.0197	0.229
	(0.0258)	(0.0189)	(0.0241)	(0.00637)	(0.0173)	(0.00810)	(0.0492)
Elder Nominee	-0.107	-0.217	0.291	0.0196	0.0903	0.0262	0.687
	(0.0250)	(0.0215)	(0.0279)	(0.00636)	(0.0227)	(0.00744)	(0.0849)
Observations	13,660	13,660	13,660	13,660	13,660	13,590	13,590
Random Household Mean	0.377	0.406	0.275	0.962	0.727	0.948	2.846
Gossip = Elder p-val	0.0382	3.32e-05	2.03e-06	0.820	0.174	0.577	6.02e-05

Notes: This table uses data from the Karnataka cell phone RCT dataset. The data consists of a individual level observations and the outcome is whether the individual nominated (as a gossip or elder, omitted is random) has the characteristic noted. Column 1 is whether the individual is SCST, column 2 is whether the primary occupation of the household is farm labor, column 3 is whether the primary income comes from land ownership, column 4 is whether the household is electrified, column 5 is whether electrification is from private purchase, column 6 is whether they own their house, column 7 is the number of rooms in the house. Standard errors (clustered at the village level) are reported in parentheses.

Appendix H. Karnataka Cellphone Experiment Payment Schedule

In the Karnataka cell phone RCT, every individual who called in and therefore was eligible for a prize was able to do the following. They simply rolled two dice. The outcome of the roll – some number between 2 and 12 – corresponded to a prize. This was independent across all participants. Every roll had some cash prize and the cell phone was worth approximately Rs. 3000. Note that if every participant rolled a 12, then every participant would win a cell phone.

Outcomes	**Pay Out**
2	25
3	50
4	75
5	100
6	125
7	150
8	175
9	200
10	225
11	250
12	Cell Phone

Figure 6. Prizes as a function of the roll of two dice in the Karnataka cell phone RCT.