Part 2: Relevance of Social Network Analysis

Jenn Larson

Vanderbilt University

June, 2019

Where we're at

Social networks—the sets of relationships interconnecting a group of people—can vary in lots of ways.

2 networks can look different overall, and 2 positions within a network can look different from each other.

What do we know about the presence of these differences and how they matter for outcomes of interest to the social sciences?

Patronage and voting

Cruz, Labonne and Querubín (2017), *AER*.

Idea: access to politicians can yield patronage benefits.

Any person would prefer that the candidate who can provide her the most benefits to win.

Patronage and voting

How measure access to politicians in the Philippines?

Marriage between families.

How acquire data on this?

Large lists of people, use naming conventions.

Patronage and voting

Cruz, Labonne and Querubín (2017), *AER*.
Observational data: construct family networks for Philippines, collect information on elections.

Show that candidates belonging to families that are more central in their villages are more likely to win elections.

Probably because people expect to receive patronage from the winners, and the family networks are instrumental for doling out the goodies.

Patronage and voting

Cruz, Labonne and Querubín (2017), *AER*.

Observational data: construct family networks for Philippines, collect information on elections.

Show that candidates belonging to families that are more central in their villages are more likely to win elections.

Probably because people expect to receive patronage from the winners, and the family networks are instrumental for doling out the goodies.

Go deep in one research design.

"Ethnic Networks," Larson and Lewis (2017), *AJPS*.

"From Chatter to Action: How Social Networks Inform and Motivate in Rural Uganda," Larson, Lewis and Rodríguez, working paper.

Ethnicity and Information Spread

Theoretical:

- Groups that can spread information widely and quickly can enforce cooperative outcomes.
  - Key to theories of economic development, governance, conflict.

Empirical:

- Ethnically diverse groups experience outcomes that are less cooperative than those of ethnically homogeneous groups.
  - Diversity empirically correlated with lower public goods provision, lower accountability, more conflict.

Inference:

- Ethnically diverse groups must be worse at spreading information widely and quickly.

The Purported Downside to Diversity

Plausible reasons:

- Common Language
- Shared understanding
- Findability
- Affinity

Assumed impact:

- Social networks within ethnic groups are dense
- Few links across ethnic groups
- $\rightarrow$ Networks in ethnically diverse areas less dense

"...assume that social sanctions and coordination are possible within groups due to the dense networks of information and mutual reciprocity that exist within groups but are not possible across groups." - Miguel and Gugerty (2005)

Conventional Wisdom

Information spreads poorly in ethnically diverse areas because social networks are less dense there.

Surprisingly little empirical investigation.

Testing the Conventional Wisdom

We find:

1. Information in fact spread worse in an ethnically diverse village than in a very similar homogeneous one.
2. Social networks are substantially *denser* in the diverse village.

Conventional Wisdom, Tested

Information spreads poorly in ethnically diverse areas ~~because social networks are less dense there~~.

We demonstrate:

3. Dense networks in diverse areas can severely hinder the spread of information if people distrust non-coethnics even a little bit.

Our Approach

Paired comparison of two villages, field experiment to measure information spread.

- Seed information with small number of randomly chosen households in two rural villages in Uganda.
- Info: event will be held in 3 days at which any adult will receive large blocks of soap in exchange for taking our survey.
- Encouraged to tell other adults.
- Team was absent from village before the seeding and between the seeding and the event.
- At event, asked about various networks and other information.
- After the event, survey others throughout the villages.

The Spread of Information

Information spread **much** more widely in Abalang (homogeneous) than in Mugana (diverse).

Outcomes of Intervention		
	Abalang	Mugana
Seeds	7	10
Proportion Heard	62%	9%

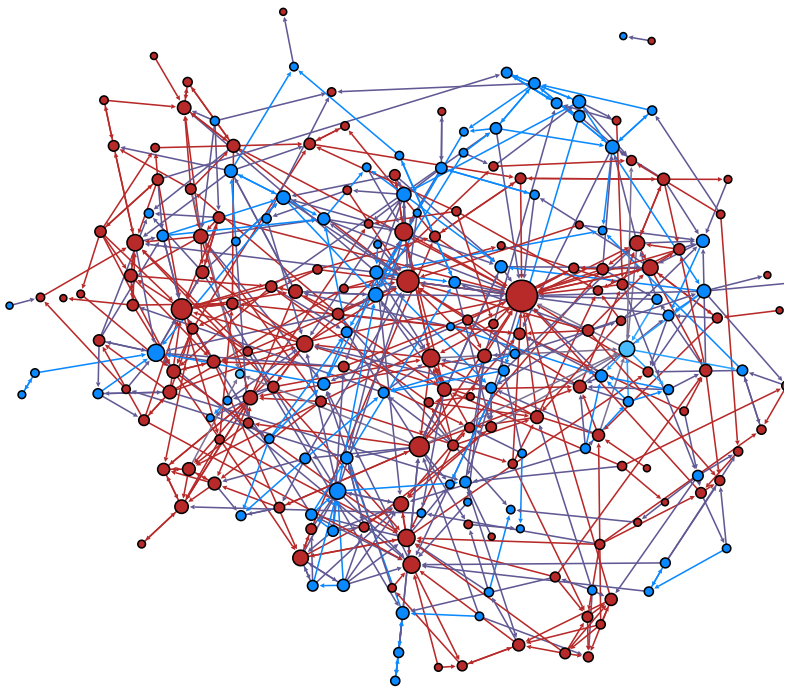
Measuring Social Networks

Ask each respondent to list up to five individuals with whom they spend time, share meals, discuss politics, discuss religion, share secrets, speak on the phone, and whom they visit.

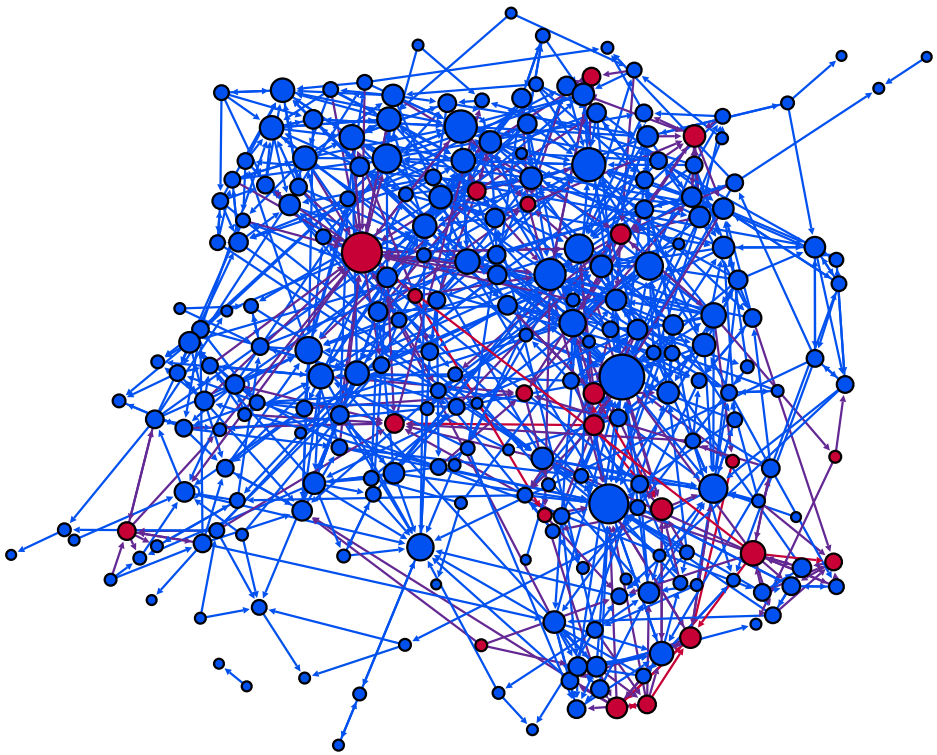
Union of these seven networks is the "social network" within each village.

Social Networks in Abalang and Mugana

Abalang



Mugana



Post-event data, nodes are respondents, red=heard, size $\propto$ degree.

Social Networks in Abalang and Mugana

Mugana's social network has many more links **and a larger proportion of the total possible links.**

	Abalang	Mugana
Node Count	216	234
Link Count	660	965
Density	**.014**	**.018**

The social network in Mugana, the ethnically diverse village, is *substantially denser* than the social network in Abalang.

Cross-Group Ties in Mugana

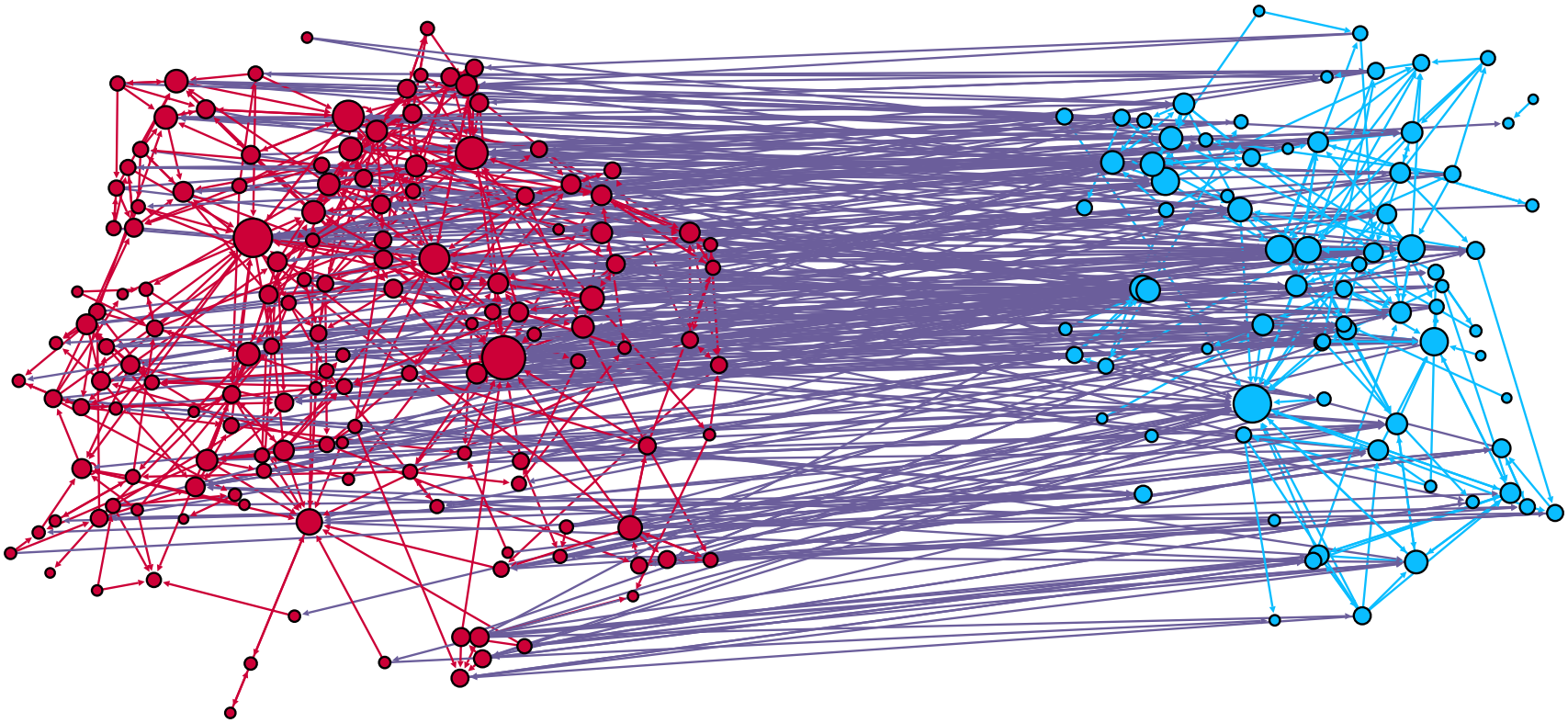


Figure: Mugana's network with the Iteso tribe on the left and Kumam on the right. 39% of social ties are between non-coethnics.

Taking Stock

News in the ethnically homogeneous village reached many more people, despite the social network being *substantially less dense* in that village.

Suggestive support for half of conventional wisdom; however also suggests the mechanism may not be density.

Conclusion

- Direct empirical evidence of information spread and ethnic networks in rural setting.
- Puzzling finding in these two villages: info spread much better in the homogeneous village, but social networks are much denser in the heterogeneous village
- Theoretical reason why density undermines info spread in heterogeneous areas in general (difference in trust toward coethncis and non-coethnics)

How does new information spread through groups, and how do people decide to act on it?

Why Care

- Lots of political outcomes hinge on people hearing AND acting on what they hear
  - Rebellion
  - Voting
  - Development Programs
  - Protest
  - Containing Health Crises
- Social Networks probably matter
- Little is known about this process

Empirically Measuring the Process of Informing and Motivating

Studies rarely designed to have all of:

1. new information
2. measures of hearing AND acting
3. direct record of how information traveled
4. verifiable action
5. precise measure of social network

Our Approach

Go deep into the process

- Seed information with small number of randomly chosen households in rural village in Uganda.
- Info: event will be held in 3 days at which adults attending will receive large blocks of soap in exchange for taking our survey.
- Encouraged to tell other adults.
- Team was absent from village before the seeding and between the seeding and the event.
- At event, asked about info sources, social ties, and other information.
- After the event, survey others in the villages, asked whether knew, social ties, and other information.

Seemingly Benign Information

While a soap event may seem boring...

Far enough outside realm of experience to trigger worry of devil worship, witchcraft

Perceived risk v. actual risk

The Village

Abalang, Uganda

- Pop: 1,400
- Predominantly peasant farmers
- Ethnically homogeneous (Ateso)
- 65% Catholic
- 13% cement-constructed houses
- Fewer than 1/3 own phone
- *Design Note: paired with nearby ethnically diverse village (Larson and Lewis 2017, AJPS)*

Response to Intervention

Attended	138
Mins Trav.	50.29 (2 - 180)
Heard, Didn't Attend	130

Measuring Networks

Ask each respondent to list up to five individuals with whom they spend time, share meals, discuss politics, discuss religion, share secrets, speak on the phone, and whom they visit.

Union of these seven networks is the "social network."

Hearers and Attenders

	Base	Hearers	Attenders	Hear v. Attend
Prop. Female	0.59	0.73*	0.79**	
Age	39.43	39.88	37.57	
Prop. Married	0.79	0.91*	0.83	*
Prop. Catholic	0.63	0.63	0.68	
WallMat	0.10	0.13	0.14	
Educ	3.34	3.33	3.23	
Prop. Unemp	0.03	0.02	0.04	
Prop. PartTime	0.42	0.47	0.46	
Prop. FullTime	0.15	0.10	0.18	*
Prop. Retired	0.39	0.42	0.32	*
Number of People	59	130	138	

*p<0.1; **p<0.05; ***p<0.01

Table: Base: neither heard nor attended. Hearers: heard but did not attend. Hearers and Attenders columns compare to Base.

Comparing Network Position

Table: Network Features of Interest

NumPeers	number of network neighbors
NumPeersPeers	avg number of network neighbors' neighbors
DistSeed	length shortest path to a seed
DistEarlyAttender	length shortest path to a Day 1 attender
AvgDist	average length shortest path to anyone
Eigen	eigenvector centrality

Calculated on full social network (union of 7) unless otherwise noted.

Hearers and Attenders: Network Comparisons

	Base	Hearers	Attenders	Hear v. Attend
NumPeers	12.51	14.67*	15.80***	
AvgNumPeersPeers	9.90	11.24**	9.52	***
PropPeersHeard	0.73	0.88**	0.92***	**
PropPeersAttended	0.14	0.18*	0.18*	
DistSeed	2.14	1.85**	2.02	*
DistEarlyAttender	2.27	1.96**	1.75***	**
AvgDist	3.92	3.78**	3.80**	
Eigen. Centrality	0.05	0.09***	0.06	**
Number of People	59	130	138	

*p<0.1; **p<0.05; ***p<0.01

Table: Base: neither heard nor attended. Hearers: heard but did not attend. Hearers and Attenders columns compare to Base.

Hearers and Attenders: Network Comparisons

	Base	Hearers	Attenders	Hear v. Attend
NumPeers	12.51	14.67*	15.80***	
AvgNumPeersPeers	9.90	11.24**	9.52	***
PropPeersHeard	0.73	0.88**	0.92***	**
PropPeersAttended	0.14	0.18*	0.18*	
DistSeed	2.14	1.85**	2.02	*
DistEarlyAttender	2.27	1.96**	1.75***	**
AvgDist	3.92	3.78**	3.80**	
Eigen. Centrality	0.05	0.09***	0.06	**
Number of People	59	130	138	

*p<0.1; **p<0.05; ***p<0.01

Table: Base: neither heard nor attended. Hearers: heard but did not attend. Hearers and Attenders columns compare to Base.

Hearers and Attenders: Network Comparisons

	Base	Hearers	Attenders	Hear v. Attend
NumPeers	12.51	14.67*	15.80***	
AvgNumPeersPeers	9.90	11.24**	9.52	***
PropPeersHeard	0.73	0.88**	0.92***	**
PropPeersAttended	0.14	0.18*	0.18*	
DistSeed	2.14	1.85**	2.02	*
DistEarlyAttender	2.27	1.96**	1.75***	**
AvgDist	3.92	3.78**	3.80**	
Eigen. Centrality	0.05	0.09***	0.06	**
Number of People	59	130	138	

*p<0.1; **p<0.05; ***p<0.01

Table: Base: neither heard nor attended. Hearers: heard but did not attend. Hearers and Attenders columns compare to Base.

Results:

Results

1) Hearing about the event followed straightforward diffusion: the more of one's peers heard, the more likely one heard.

2) Attending did not: even conditional on hearing, having more peers attend does *not* make one more likely to attend.

Suggests that if the network matters for attendance, it is not via simple social endorsement.

Credibility

Two options:

1. Seeds
2. Early Attenders

Seeds received the information first, were given the official information, and were provided an information sheet.

Early Attenders saw the event with their own eyes, came back with soap.

Social Proximity to Early Attenders

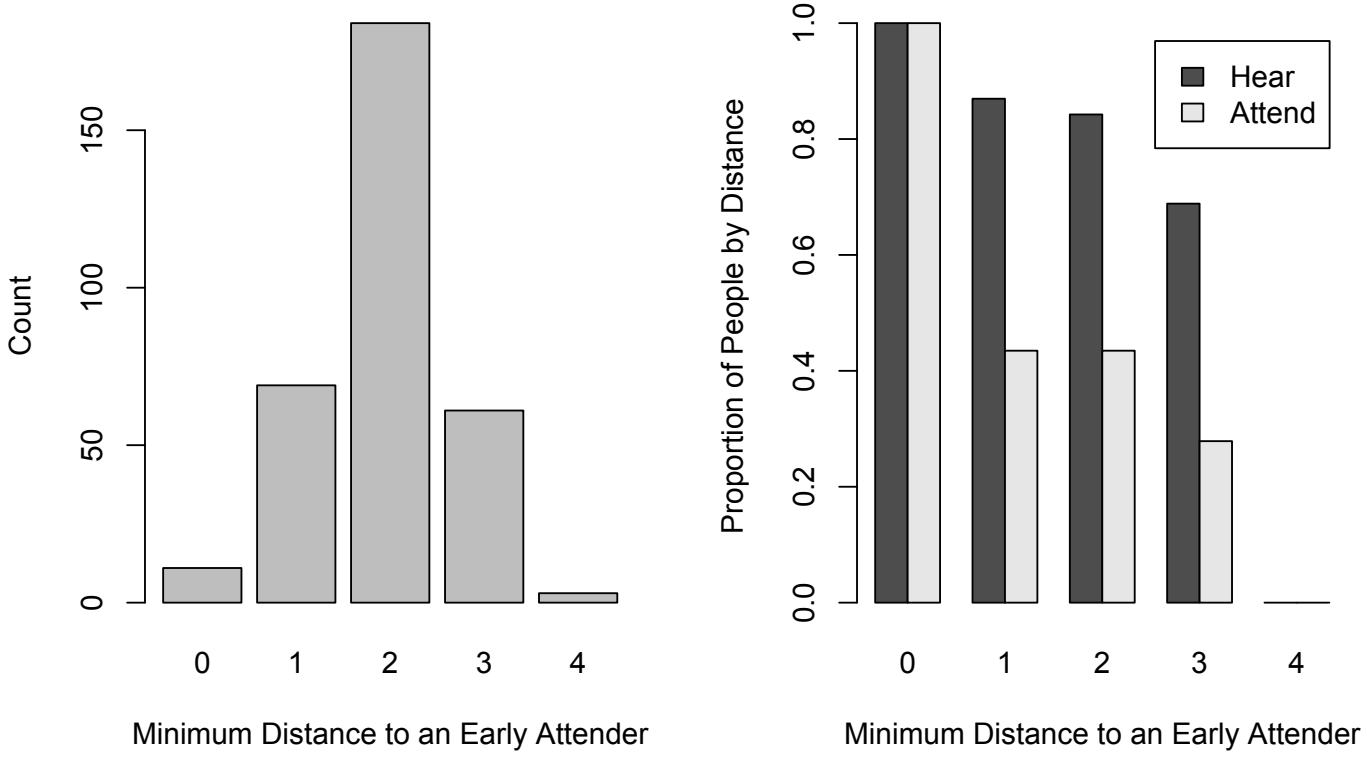


Figure: Proportion of those in the sample at different distances to an early attender who heard and attended. Fewer of those farther from an early attender attended.

Social Proximity to Seeds

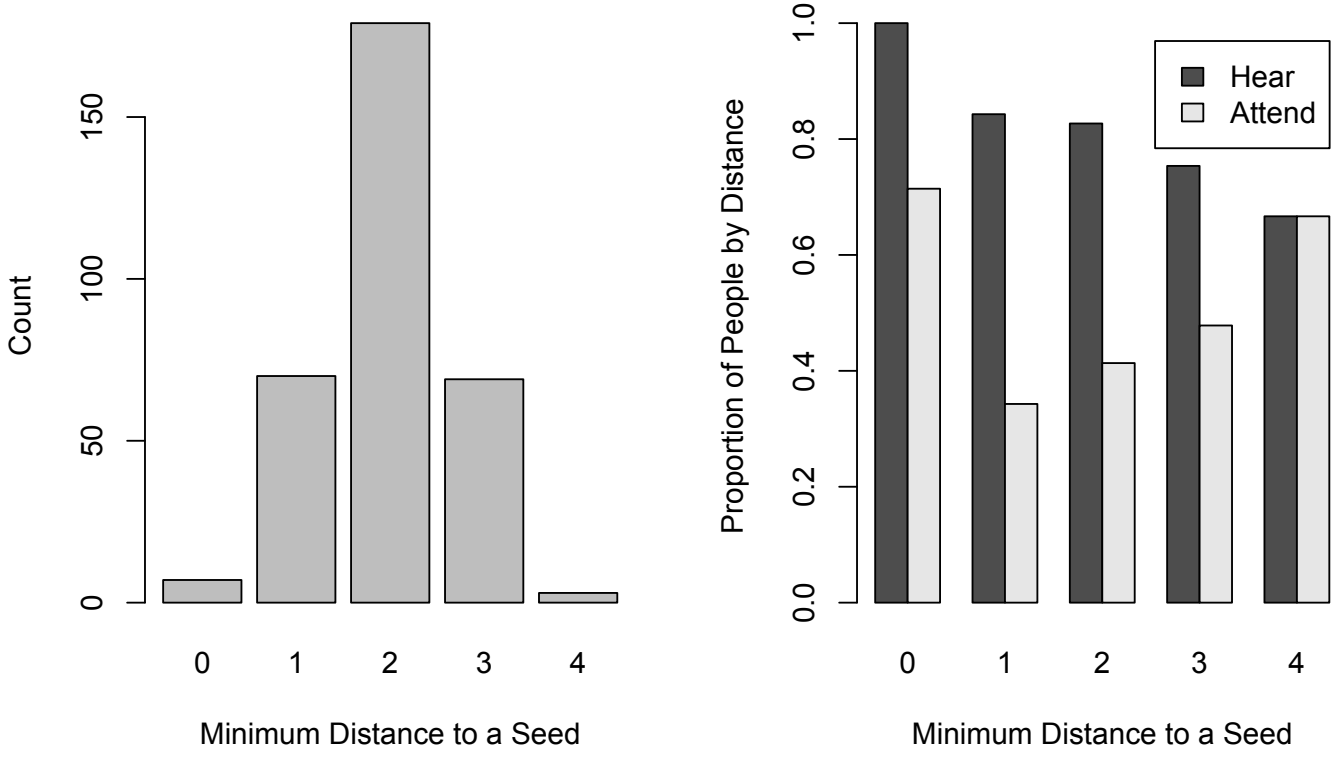


Figure: Proportion of those in the sample at different distances to a seed who heard and attended. *More* of those farther from the seeds attended.

Who Are the Early Attenders?

	Early Attenders	Later Attenders
Female	0.82	0.79
Age	43.27	37.07
Married	0.91	0.83
Catholic	0.64	0.69
WallMat	0.18	0.13
Educ	3.09	3.24
Unemp	0.09	0.04
PartTime	0.45	0.46
FullTime	0.09	0.18
Retired	0.36	0.32
Number of People	11	127

*p<0.1; **p<0.05; ***p<0.01

Who Are the Early Attenders?

	Early Attenders	Later Attenders
NumPeers	17.55	15.65
AvgNumPeersPeers	11.81	9.32**
PropPeersHear	0.94	0.91
PropPeersAttend	0.32	0.17**
DistSeed	1.27	2.09**
DistEarlyAttend	0.00	1.90***
DistOtherEarlyAttend	1.09	1.90***
AvgDist	3.69	3.81
Eigen	0.09	0.06
Number of People	11	127

*p<0.1; **p<0.05; ***p<0.01

Who Are the Early Attenders?

	Early Attenders	Later Attenders
NumPeers	17.55	15.65
AvgNumPeersPeers	11.81	9.32**
PropPeersHear	0.94	0.91
PropPeersAttend	0.32	0.17**
DistSeed	1.27	2.09**
DistEarlyAttend	0.00	1.90***
DistOtherEarlyAttend	1.09	1.90***
AvgDist	3.69	3.81
Eigen	0.09	0.06
Number of People	11	127

*p<0.1; **p<0.05; ***p<0.01

Results

3) Social proximity to an early attender makes one more likely to attend and matters more than social proximity to a seed

Disaggregating Network Type

Table: Attending $\sim$ PropPeersAttend by Network Type

Tie Type	p-val	Obs.
Time		263
Phone	*	263
Politics		183
Religion		251
Meal	**	278
Visit	**	229
Secrets		226

Consistent with contagion along Meal and Visit ties, i.e. an endorsement effect among these peers.

Results

4) Endorsement by especially intimate peers may matter for attendance

Explaining Outcomes, Including Controls

	Hear
PropPeersHear	0.502***
	(0.116)
NumPeers	0.002
	(0.003)
Eigen	0.720***
	(0.275)
DistEarlyAttend	−0.065**
	(0.030)
Female	0.143**
	(0.061)
Age	0.0002
	(0.001)
Catholic	0.033
	(0.039)
Educ	0.003
	(0.013)
Married	0.052
	(0.068)
WallMat	0.077**
	(0.036)
Adj. R-Squared	0.23
Observations	310

*p<0.1; **p<0.05; ***p<0.01

	Attend
PropPeersAttend	−0.259
	(0.294)
NumPeers	0.010*
	(0.006)
Eigen	−1.321**
	(0.608)
DistEarlyAttend	−0.123**
	(0.053)
Female	−0.009
	(0.089)
Age	−0.002
	(0.002)
Catholic	0.016
	(0.071)
Educ	−0.037
	(0.025)
Married	−0.159*
	(0.093)
WallMat	0.055
	(0.092)
Adj. R-Squared	0.069
Observations	256

*p<0.1; **p<0.05; ***p<0.01

Explaining Outcomes, Including Controls

	Hear
PropPeersHear	0.502***
	(0.116)
NumPeers	0.002
	(0.003)
Eigen	0.720***
	(0.275)
DistEarlyAttend	−0.065**
	(0.030)
Female	0.143**
	(0.061)
Age	0.0002
	(0.001)
Catholic	0.033
	(0.039)
Educ	0.003
	(0.013)
Married	0.052
	(0.068)
WallMat	0.077**
	(0.036)
Adj. R-Squared	0.23
Observations	310

*p<0.1; **p<0.05; ***p<0.01

	Attend
PropPeersAttend	−0.259
	(0.294)
NumPeers	0.010*
	(0.006)
Eigen	−1.321**
	(0.608)
DistEarlyAttend	−0.123**
	(0.053)
Female	−0.009
	(0.089)
Age	−0.002
	(0.002)
Catholic	0.016
	(0.071)
Educ	−0.037
	(0.025)
Married	−0.159*
	(0.093)
WallMat	0.055
	(0.092)
Adj. R-Squared	0.069
Observations	256

*p<0.1; **p<0.05; ***p<0.01

Explaining Outcomes, Including Controls

	Hear
PropPeersHear	0.502***
	(0.116)
NumPeers	0.002
	(0.003)
Eigen	0.720***
	(0.275)
DistEarlyAttend	−0.065**
	(0.030)
Female	0.143**
	(0.061)
Age	0.0002
	(0.001)
Catholic	0.033
	(0.039)
Educ	0.003
	(0.013)
Married	0.052
	(0.068)
WallMat	0.077**
	(0.036)
Adj. R-Squared	0.23
Observations	310

*p<0.1; **p<0.05; ***p<0.01

	Attend
PropPeersAttend	−0.259
	(0.294)
NumPeers	0.010*
	(0.006)
Eigen	−1.321**
	(0.608)
DistEarlyAttend	−0.123**
	(0.053)
Female	−0.009
	(0.089)
Age	−0.002
	(0.002)
Catholic	0.016
	(0.071)
Educ	−0.037
	(0.025)
Married	−0.159*
	(0.093)
WallMat	0.055
	(0.092)
Adj. R-Squared	0.069
Observations	256

*p<0.1; **p<0.05; ***p<0.01

Explaining Outcomes, Including Controls

	Hear
PropPeersHear	0.502***
	(0.116)
NumPeers	0.002
	(0.003)
Eigen	0.720***
	(0.275)
DistEarlyAttend	−0.065**
	(0.030)
Female	0.143**
	(0.061)
Age	0.0002
	(0.001)
Catholic	0.033
	(0.039)
Educ	0.003
	(0.013)
Married	0.052
	(0.068)
WallMat	0.077**
	(0.036)
Adj. R-Squared	0.23
Observations	310

*p<0.1; **p<0.05; ***p<0.01

	Attend
PropPeersAttend	−0.259
	(0.294)
NumPeers	0.010*
	(0.006)
Eigen	−1.321**
	(0.608)
DistEarlyAttend	−0.123**
	(0.053)
Female	−0.009
	(0.089)
Age	−0.002
	(0.002)
Catholic	0.016
	(0.071)
Educ	−0.037
	(0.025)
Married	−0.159*
	(0.093)
WallMat	0.055
	(0.092)
Adj. R-Squared	0.069
Observations	256

*p<0.1; **p<0.05; ***p<0.01

Results

5) Those more eigenvector central are more likely to hear; conditional on hearing, those *less* eigenvector central are more likely to attend.

Who Are the Eigenvector Central?

	Bottom Eigen Central	Top Eigen Central
Female	0.75	0.30**
Age	42.58	39.89
Catholic	0.95	0.32***
WallMat	0.11	0.10
Married	0.79	0.89
Educ	3.11	3.70
Unemp	0.00	0.10
PartTime	0.50	0.25
FullTime	0.11	0.10
Retired	0.39	0.55
NumPeers	7.20	22.70***
NumPeersPeers	6.71	13.05***
DistSeed	2.75	1.35***
DistEarlyAttend	2.75	1.80***
AvgDist	4.49	3.48***
PropPeersHear	0.90	0.90
PropPeersAttend	0.18	0.17
Eigen	0.004	0.349***
Hear	0.80	0.90
Attend	0.55	0.35
AttendGivenHeard	.69	.39**
Number	20	20

Table: Comparison of the twenty villagers with the highest eigenvector centrality and the twenty with the lowest eigenvector centrality in the social network.

So many questions left unanswered

- Would all information work this way?
- Does this generalize to other places?
- Why do the eigenvector central not show up?

Now we are familiar with examples of networks studies.

Next up: how can we design our own study to gather network data in the field?

Break!