MULTIPLEXING IN NETWORKS AND DIFFUSION

ARUN G. CHANDRASEKHAR[‡], VASU CHAUDHARY[§], BENJAMIN GOLUB[§], AND MATTHEW O. JACKSON[†]

Abstract. A network is called multiplexed if a pair of people (nodes) can be linked through multiple types of relationships. For instance, they may borrow from one set of people and seek advice from a (possibly overlapping but) different set of people. We make three contributions to the study of multiplexing. First, we document facts about multiplexing in two different datasets. Network layers such as socializing, advising, helping, and financial ties exhibit significant correlations but also distinct patterns. Further, subcaste and geography—often used proxies for social networks—have little correlation with other layers. Second, using data from a randomized controlled trial of information diffusion, we show that advice links are more predictive of diffusion than other types of links; but also that the combination of layers is significantly more predictive than any single one. We also show that villages with higher multiplexing (greater overlap among layers) experience less diffusion of information in the RCT. Third, we develop a model of diffusion with many network layers, and show that multiplexing impedes simple contagions like disease or simple information spread but, depending on the virality of the process, may enhance or impede complex contagions that require reinforcement from multiple interactions like the spread of a new technology. Finally, we document heterogeneities in multiplexing by gender and connectedness.

Keywords: Networks, Social Networks, Multiplex, Multi-Layer, Diffusion, Contagion, Complex Contagion

JEL codes: D85, D13, L14, O12, Z13

1. Introduction

People have many different types of relationships—for example, collaborating with colleagues at work, relying on friends and acquaintances for assistance and advice, and engaging in the borrowing and lending of money or goods with family and friends. A given pair of people can have multiple such relationships. For example, people have

Date: August 14, 2024.

[‡]Department of Economics, Stanford University; NBER; J-PAL.

[†]Department of Economics, Stanford University; Santa Fe Institute.

[§]Department of Economics, Northwestern University.

We thank Paul Goldsmith-Pinkham for helpful conversations. Research was supported by NSF grants SES-1629328 and SES-2018554. Chandrasekhar acknowledges the Alfred P. Sloan Foundation for financial support.

in-person as well as online interactions; and some pairs are linked in both the offline and online networks, but the overlap between these networks is far from complete (Boccaletti et al., 2014). College students' partners in social activities overlap with, but also differ from, the people to whom they turn in times of stress or for academic collaboration (Morelli et al., 2017; Jackson et al., 2024).

This coexistence of distinct types of relationships among the same population is referred to as *multiplexing* (see, e.g., Kivela et al. (2014)), and the fact that different types of relationships are interdependent has been discussed since Simmel (1908). Although numerous case studies have examined multiple relationships, there are many basic open questions about the empirical patterns of multiplexing in social and economic relationships, as well as the relationship between multiplexing and outcomes of interest, such as information diffusion. In this paper, we make three contributions.

First, in Section 2, we present an unsupervised statistical perspective on the correlation structure across the network layers. We document that there is significant correlation between different network layers (types of relationships), but also reveal meaningful differences in their patterns. Without properly accounting for the multigraph nature of relationships a researcher may arrive at misleading conclusions about peer effects and influence. We show that layers of informational relationships, financial relationships, social relationships, among others, are strongly correlated in a sample of 143 villages in Karnataka, India, comprising nearly thirty thousand households (Banerjee et al., 2013, 2019, 2024b). Even though the layers are correlated, we also show that they exhibit distinct patterns and differ in density and other network statistics. We also show that proxies for social networks used in the peer effects literature, such as geographic neighbors or co-ethnicity (here the analogy is being members of the same *jati*, or subcaste), are nearly orthogonal to the other layers suggesting that those estimates may be problematic.

Second, we adopt a supervised statistical perspective to show that this distinction between layers matters substantively when studying economic outcomes. While one might expect little distinction between network layers in predicting the outcome of interest, we find significant distinctions and a nuanced overall picture. In the context of a randomized controlled trial of information diffusion, we show that some layers are more predictive than others of diffusion, and moreover that a combination of layers is more predictive than any given layer. We also show that a combination does better than an overall union or intersection of layers, so the layers are not simply noisy observations of the same underlying relationships.

Interestingly, the proxy of jati is least predictive of diffusion, and it is not a good substitute for the actual relationships. Nonetheless, when coupled with the other layers it turns out to significantly improve the prediction of information diffusion. Thus, although the jati layer is not a good substitute for network data, it is a good complement to it.[1]

We also show that villages that are more multiplexed (have more strongly correlated layers) experience less diffusion, which also suggests that the level of correlation between layers has an impact on how information diffuses. This fact especially sets up our theoretical analysis.

Third, in Section 3 we develop new theory of how multiplexing influences diffusion and contagion. We show that multiplexing impedes standard diffusion and contagion in which a person may be infected/informed by any single infected other. In particular, we show that comparing two individuals with the same number of relationships in every layer—with one having relationships that are more multiplexed (overlap more in a dominance sense)—the more multiplexed individual is less likely to become infected for any given probability of neighbor infection We then use this lemma to show that in a standard SIS (Susceptible-Infected-Susceptible) model, the steady-state infection rate decreases as individuals become more multiplexed. We also develop a theory of how multiplexing impacts complex diffusion, in which people only become infected or adopt a new behavior or norm if they experience sufficiently many interactions with infected others. Here we show that multiplexing either enhances or impedes diffusion, depending on how viral the process is.

We close the paper in Section 4 with some observations about how multiplexing depends on characteristics. We find that women's networks are significantly more multiplexed than men's networks, and also that multiplexing is positively correlated with a poverty measure and negatively correlated with how many connections a person has. Given the theory and experiments showing how multiplexing can impede information diffusion, differences in multiplexing can lead to differences in access to information across gender and wealth levels.

The literature on multiplex networks has begun to grow in the last decade (Contractor et al., 2011; Boccaletti et al., 2014; Kivela et al., 2014; Dickison et al., 2016; Bianconi, 2018), even though the recognition of the importance of the fact that people are involved in different types of relationships dates to some of the original network analysis (e.g.,

[1]As we discuss below, it drastically over-predicts links within jati, and underpredicts them across jati. One conjecture as to why the jati layer helps in predicting diffusion, is that patterns of information passing on the network are still related to jati.

Simmel (1908)), and various instances of the fact that different layers can have different roles have been analyzed over time (Wasserman and Faust, 1994).

More recently, various analyses have shown that tracking the interplay and distinguishing between different networks can be important in understanding cooperative behavior (Atkisson et al., 2019; Cheng et al., 2021) as well as understanding and targeting people's play in network games (Walsh, 2019; Zenou and Zhou, 2024).

The contributions to the multiplex literature that we make here are threefold.

First, we provide some of the first detailed statistical analysis of how the layers relate to each other.

Second, we show how different layers, as the level of correlation, between layers predict diffusion. This suggests that unidimensional theories of diffusion and contagion need to be rethought.

Third, we provide the first theoretical analysis of how correlation between layers impacts diffusion, which provides a basis for answering the above question. Although there has been some theoretical analysis of simple (Hu et al., 2013; Larson and Rodriguez, 2023) and complex (Yağan and Gligor, 2012; Zhu et al., 2019; Kobayashi and Onaga, 2023) diffusion on multiplexed networks, the analysis has focused on independently distributed layers. Such diffusion is a more direct extension of diffusion on one layer and the proofs in that literature leverage that fact. Here, we analyze how changes in the extent to which layers coincide with each other impacts diffusion. In addition, the new model that we develop also allows for correlated interactions across layers. Thus, we allow for correlation both in the contact across layers, and the structure of the layers themselves, neither of which appear in prior results.

Our findings on the impact of multiplexing on diffusion can help inform a nascent and important literature on the incentives to form multiplexed networks (Billand et al., 2023; San Román, 2024). For example, a series of empirical studies in rural developing economies emphasize the role social networks play in risk-sharing arrangements (Townsend, 1994; Fafchamps and Gubert, 2007; Ambrus et al., 2014) This raises a fundamental question: If individuals primarily organize their relationships around risk-sharing and multiplex other relationships on top of the risk-sharing relationships, how might these structures affect the diffusion of new information or technologies? Our empirics and theory both provide answers.

2. Multiplexing in the Data

2.1. **Two Data Sets.** We study two different samples of village network data totaling 143 villages, both from the state of Karnataka, India, covering a total population of nearly thirty thousand households.

2.1.1. *The Microfinance Village Sample.* The first dataset that we use—and refer to as the "Microfinance Village Sample"—comes from the Wave II network data collected in a study of 75 villages (Banerjee et al., 2013, 2024b). In 2012, we obtained a complete census of the 16,476 households across the 75 villages. We then went to 89.14% of the households and collected detailed socio-economic network data, described below. (This means that we obtained information on 98.8% of the links[2].)

We collected information on various types of interactions for each respondent, spanning social, financial, informational, kinship, and religious networks. We asked respondents about various types of relationships, each listed below with an abbreviated label that we use from now on:

(1) social: to whose house the respondent goes to and who comes to their home, as well as which close relatives live outside of their household.
(2) kerorice: from whom does the respondent borrow kerosene/rice and to whom does the respondent lend these goods;
(3) advice: to whom does the respondent give information/advice;
(4) decision help: to whom does the respondent go to get help with an important decision;
(5) money: if the respondent borrow suddenly needed 50 rupees for a day, to whom would they turn, and whom would come to them with such a request;
(6) temple: if the respondent goes to a temple, church or mosque, who might accompany them;
(7) medic: if the respondent had a medical emergency alone at home, whom would they ask for help in getting to a hospital.

In addition to these layers, we have information on the jati (subcaste) as well as GPS coordinates for each household. These can be used to construct networks by considering pairs linked if they are from the same jati or live closer than some threshold. These sorts of variables have been used in social science as proxies for social networks (Sacerdote, 2001; Fafchamps and Gubert, 2007; Munshi and Rosenzweig, 2009).

[2]Given our focus on undirected graphs, we elicit a link as long as at least one of the two households on either end is sampled. With 89.14% of the households being sampled, for any two arbitrary nodes i and j, $\mathrm{P}(i \text{ or } j \text{ in sample}) = 1 - (1 - 0.89)^2 = 0.9879$

2.1.2. *The Diffusion RCT Sample.* The second dataset that we use is what we call the "RCT villages." It contains multiple network layers in another 68 villages collected by (Banerjee et al., 2019).

The network data was collected in a manner similar to that of the Microfinance Village Sample. In these surveys we collected information about the following layers:

(1) social: to whose house the respondent goes to and who comes to their home to socialize;
(2) kerorice: from whom does the respondent borrow kerosene/rice or small amounts of money and to whom does the respondent lend these goods;
(3) advice: to whom does the respondent give information/advice;
(4) decision: to whom does the respondent go to get help with an important decision.

While we also have information on jati in the RCT villages, we do not have GPS data for this sample.

In addition, we conducted a randomized controlled trial (RCT) on diffusion in these villages, which is the subject of Banerjee et al. (2019). Data from this RCT allow us to take advantage of cleanly identified estimates of diffusion to assess how diffusion varies with various aspects of multiplexed networks. Specifically, in each village we randomly seeded either 3 or 5 individuals (determined uniformly at random) with information about a promotion. Villagers had a (non-rivalrous) chance to win either cash prizes or a mobile phone if they called in to register for the promotion.[3] Registered callers were visited a few weeks later and received a reward.[4] Thus, we experimentally induced the diffusion of a non-rivalrous, valuable piece of information. The outcome variable of interest is the number of households that registered. There was exogenously randomized variation in the position of the random seeds in the network, and more central seeds caused larger diffusions. We use our data on multiplexing to examine how diffusion depends on which network layer is used to measure seed position, and how multiplexed the network is.

2.2. **The Layers and Multiplexing Patterns.**

We begin with some descriptive statistics on the various network layers.

[3]In particular, they had to dial the provided promotional number and leave a "missed call." This was a call that we registered but did not answer and was free for the participant to make, which was a standard technique for registration at the time.

[4]The individual rolled a pair of dice and received INR 25 $\times$ the number rolled. This yielded cash prizes of amounts ranging from INR 50 (for a 2) to INR 275 (for an 11). A roll of 12 was rewarded with a cell phone worth INR 3000. The expected value of the prize was INR 255, which was more than half of a day's wage in the area.

TABLE 1. Descriptive Statistics

Network	degree	degree S.D.	density	triangles	clustering
Microfinance villages					
social	15.296	7.841	0.079	2635.040	0.252
kerorice	7.029	3.834	0.037	594.160	0.259
advice	6.158	3.835	0.032	299.120	0.168
decision	6.553	4.309	0.034	356.040	0.169
money	8.512	5.036	0.044	681.960	0.193
temple	1.709	1.899	0.009	52.040	0.175
medic	6.530	3.911	0.034	369.400	0.188
union link	75.428	32.542	0.368	314121.027	0.862
intersect link	0.576	0.883	0.003	7.000	0.203
jati	68.291	34.293	0.332	310150.907	1.000
RCT villages					
social	5.711	3.626	0.031	251.271	0.185
kerorice	4.910	3.235	0.027	176.557	0.174
advice	4.197	3.091	0.023	124.100	0.161
decision	4.206	3.675	0.023	125.571	0.145
union link	55.756	27.861	0.296	150771.400	0.913
intersect link	1.812	1.829	0.010	38.871	0.229
jati	52.633	28.599	0.279	150117.500	1.000

A link is present in a given layer if either household named the other household in one of the questions in that category (e.g., we code a kerorice link if either household said that the other household borrows kerosene or rice from them or lends it to them). In terms of notation, we define a multi-layered, undirected[5] network for each village v, for layer $\ell = 1, \ldots, L$, with $g^{\ell}_{ij,v} = 1$ if either household i or j said they had a relationship of type ℓ. To the survey layers, we add another layer where i and j are linked if they belong to the same jati. For the Microfinance Village sample, where GPS data are available, we construct a weighted graph where ij entry is the geographic distance between the two households.

The *union* layer has a link present if a link is present in any layer. The *intersection* layer has a link present if it is present in all layers.[6]

[5]One can also define directed networks from our data, which we comment on at several places below. Directed links open some important but tangential questions for multiplexing that we prefer to leave for further research.

[6]Both of these definitions include the jati layer but exclude geography, since we are able to define the geographic layer for only one of our datasets and it is a weighted network in any case. We make these definitions to maintain consistency of the meaning of the union and intersection layers across the two data sets.

2.2.1. *Descriptive Statistics.* Our first look at the data focuses on basic descriptive statistics, presented in Table 1. The different layers have (significantly) different patterns of connection. For example, the social layers are denser than the other layers and have among the highest levels of clustering, and we see higher variance of node degrees in the decision layer than in other comparable layers (e.g., advice).[7]

We also see that the microfinance villages and RCT villages differ from each other.

In addition, we see that the jati layer, where nodes are linked if they are of the same jati, has by far the highest degree. This foreshadows the fact that jati ends up being a poor proxy for the other types of relationship. It ends up being too dense, too clustered, and too homophilous to match up with the other layers.

2.2.2. *Correlations among Layers.* Next, we examine the correlation between layers, pictured in Figure 1.

Figure 1 shows several things. First, there are consistently high correlations between layers in both data sets—above 0.5 for most layer pairs. Second, the exceptions are the distance, jati, and temple layers. The jati and distance layers are almost uncorrelated with the other layers,[8,9] while the temple layer has an intermediate level of correlation with others. Third, the layers are more highly correlated in the RCT compared to the microfinance villages.

2.2.3. *Principal Component Analyses.* Our third look at the structure of the networks consists of principal component analyses. We do this in two passes.

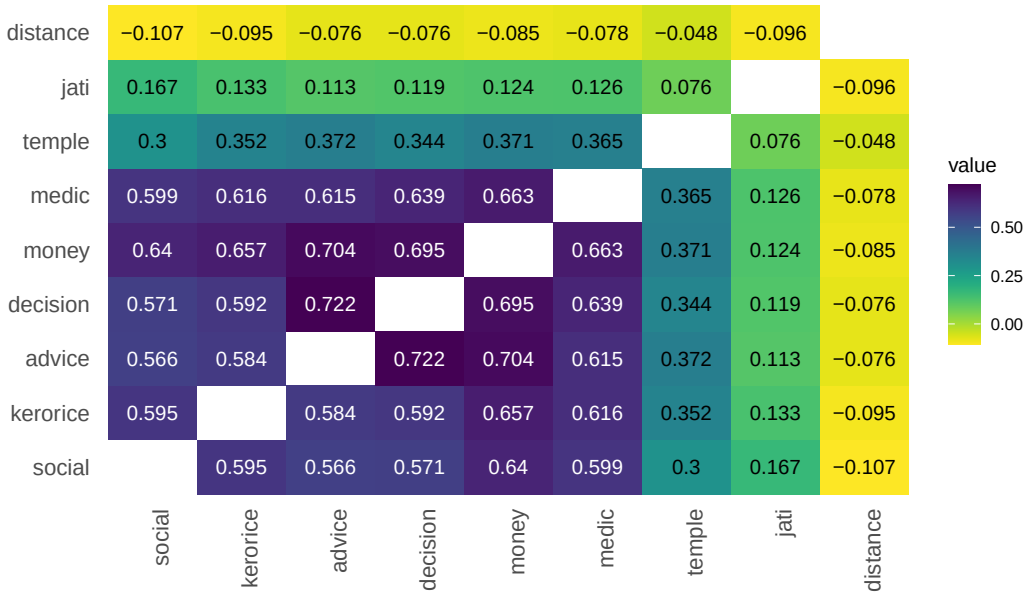
First, we perform a principal component analysis with all of the layers (excluding *union* and *intersection*). Specifically, we consider each pair of households (in a given village) an observation, yielding $\sum_v \binom{n_v}{2}$ observations, where n_v is the number of households in the village, and the number of dimensions is the number of layers L in the given sample.

As we see in Figure 2, when we include all of the layers, the first principal component lines up with most of the relationship layers, picking up almost half (48.7%) of the variation in the microfinance villages and more than two-thirds (72.1%) in the RCT villages. Panels C and D plot the coordinates of the first and second component entries for each link type. Interestingly, jati largely aligns with the second dimension, as one

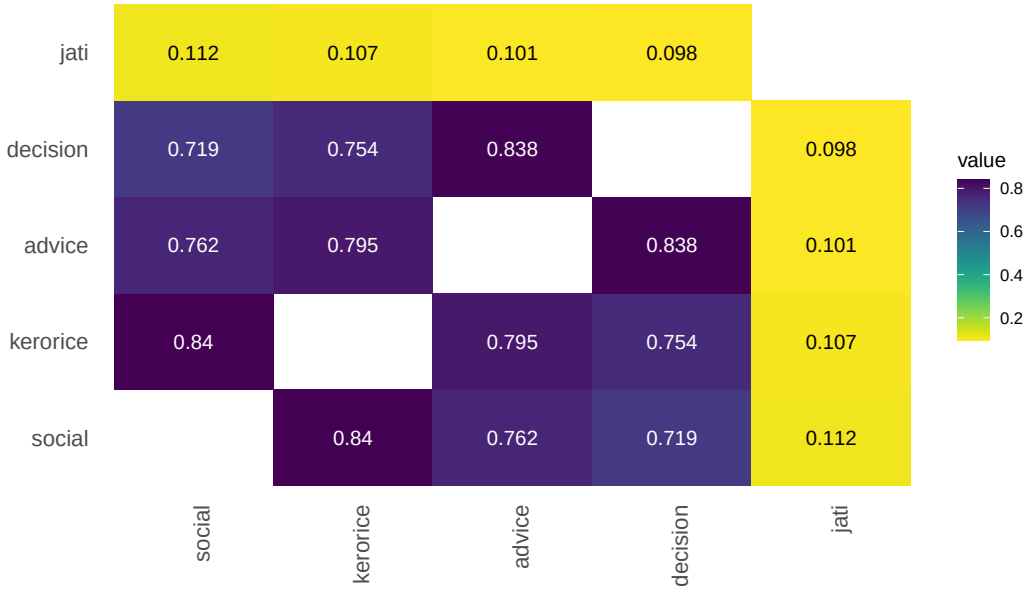
[7]In the RCT villages the social layer is significantly more dense than kerorice, advice or decision layers (p-values 0.009, 0.000 and 0.000 respectively). The advice layer has significantly less variance (p-value 0.0006) and more clustering (p-value 0.03) than the decision layer. Similar patterns hold in the Microfinance villages.

[8]This does not mean, for instance, that there is not substantial jati-based homophily in these data. The low correlation comes from the fact that the jati layer dramatically over-predicts relationships compared to other layers, so it has many 1's where there are 0's in the other layers.

[9]Distance is higher when people live far from each other, and are thus less likely to be linked all else held equal, and hence the negative signs.



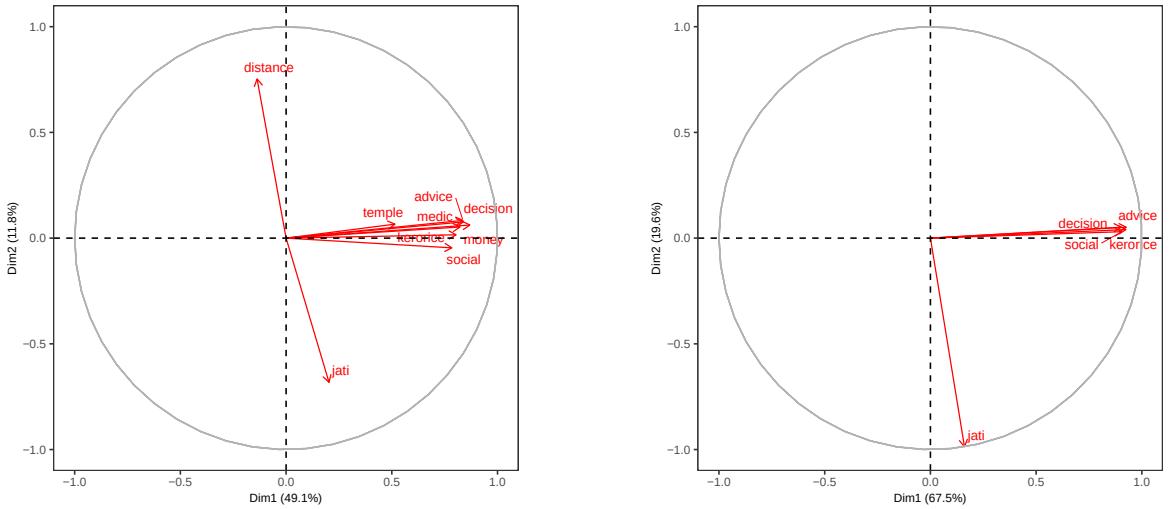
(A) Microfinance Villages



(B) RCT Villages

FIGURE 1. Correlation Heatmaps

would expect given its relatively low correlation with the other layers. Geography is nearly opposite of jati, which reflects the fact that people from the same jati live in close proximity. The full table is provided in Appendix Tables S1 and S2).



(A) Principal Components: Microfinance Villages

(B) Principal Components: Diffusion RCT

FIGURE 2. Principal Component Analysis with All Layers

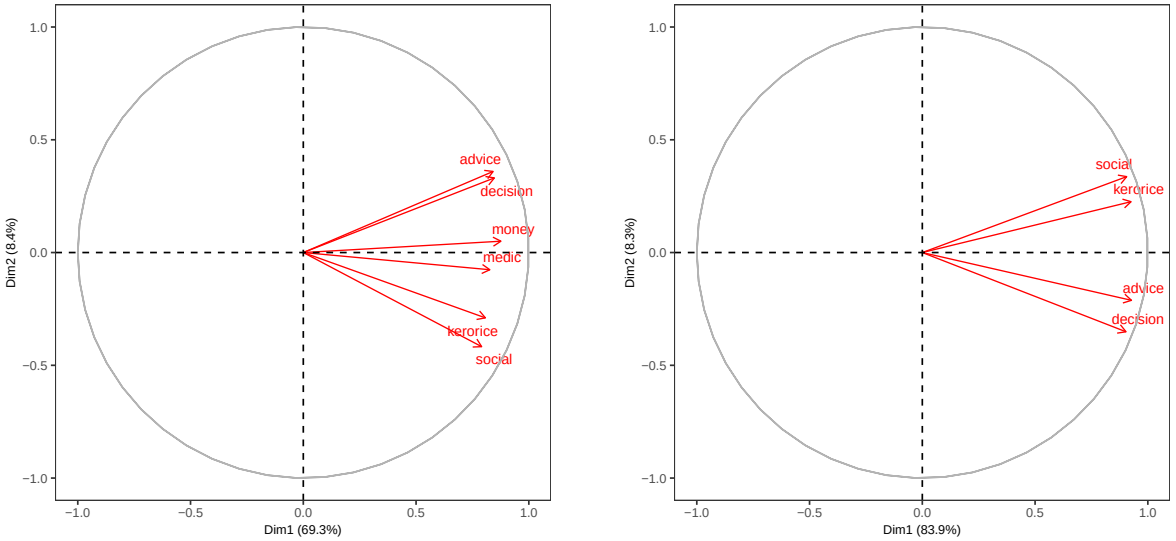
Next, in Figure 3 we redo the analysis when we drop the least correlated dimensions of jati, geography and temple.[10]

Figure 3 exhibits three distinct pairings in the microfinance villages (advice-decision, money-medic, kerorice-social) and two in the RCT villages (advice-decision, kerorice-social), with the first component now explaining 70% and 83% percent of the variance across the two samples, respectively.

Building the Backbone. To capture the correlation structure of the network layers we use the principal component analysis to construct a layer from the multigraph that we call the *backbone*. The backbone layer is built using the first K principal components constructed above. The optimal K is determined by the ladle plot constructed from a combination of the scree plot and bootstrap eigenvector plot (see Luo and Li (2016) for details).[11]

[10]We also redo the analysis just dropping jati and geography and keeping temple in Supplemental Appendix Figure S2. Temple is sparse and essentially orthogonal to the other dimensions.

[11]To select the optimal number of principal components the literature usually relied on a cutoff based on patterns of either decreasing eigenvalues or increasing variability of eigenvectors. Luo and Li (2016) combine these two approaches to better estimate the optimal K. They propose a new estimator, called "ladle estimator" which minimizes an objective function that incorporates both the magnitude of eigenvalues and the bootstrap variability of eigenvectors. This approach exploits the pattern that when eigenvalues are close together, their corresponding eigenvectors tend to vary greatly, and when eigenvalues are far apart, the eigenvector variability tends to be small. By leveraging both sources of information, the ladle estimator can more precisely determine the rank of the matrix, and thus the optimal number of components to retain.



(A) Principal Components: Microfinance Villages

(B) Principal Components: Diffusion RCT

FIGURE 3. Principal Component Analysis Excluding Jati, Geography, and Temple

For a pair ij in village v we compute the weighted sum of its projections on the first K principal components as

$$Z_{ij,v} = \sum_{k=1}^{K} w_k \cdot \left(\sum_{\ell=1}^{L} g_{ij,v}^{\ell} \cdot e_{k\ell} \right).$$

In this formula, e_k is the eigenvector associated with the k^{th} principal component, and the weights w_k are determined by relative magnitudes of the eigenvalues associated with each component:

$$w_k := \lambda_k / \sum_{j=1}^{K} \lambda_j.$$

For each village v, we determine a cutoff score z_v based on the average empirical density of the network layers d_v such that $\mathbb{P}(Z_{ij,v} > z_v) = d_v$. Finally, we generate the backbone graph $g^{backbone}$ as

$$g_{ij,v}^{\text{backbone}} = \mathbf{1}\{Z_{ij,v} > z_v\}.$$

2.3. **Determinants of Diffusion.**

The empirical analysis shows that multiplexed networks in our data are rich and embed important information that would be lost by collapsing them into a single summary measure. A next natural question is how this distinction between layers matters for an outcome of interest. We focus here on diffusion in the RCT villages.

We proceed as follows. Based on prior work, we expect that more central seeds should lead to greater diffusion (Banerjee et al., 2013). Those papers defined a single network by using the union network and computed diffusion centrality based on that. However, given the rich multiplex data in the multiplexing data, we note that the seeds' diffusion centrality differs across layers. Thus, we examine which layer is the most predictive of diffusion in the RCT if we compute diffusion centrality based on it alone.

We use a particular diffusion centrality measure developed in that paper and further studied in Banerjee et al. (2019). In particular, the *diffusion centrality* of a node j in layer ℓ in village v, $DC^{\ell}_{j,v}$ is defined by:

$$DC^{\ell}_{j,v} := \left[\sum_{t}^{T}(qg^{\ell}_{v})^{t} \cdot 1\right]_{j},$$

where T is the number of rounds of communication and q is the probability of transmission in each period across any given link. Following Banerjee et al. (2019), for village v and network layer ℓ, we set $T = \text{diameter}(g^{\ell}_{v})$, and $q = 1/\lambda^{\ell}_{v}$, where λ^{ℓ}_{v} is the largest eigenvalue associated with g^{ℓ}_{v}. We calculate the diffusion centrality of the seed set of village v, S_v, for layer ℓ by

$$DC^{\ell}_{v} := \sum_{j \in S_v} DC^{\ell}_{j,v}.$$

We then can calculate how diffusion varies with the diffusion centrality of the randomly assigned seed set under a layer ℓ by regressing

$$y_v = \alpha^{\ell} + \beta^{\ell} \cdot DC^{\ell}_{v} + X_v\Gamma^{\ell} + \epsilon_{v,\ell} \tag{2.1}$$

where y_v is the number of calls received from village v, and X_v includes controls for number of households, its second and third powers, and number of seeds assigned in that village. We standardize all of the regressors.

We can see how differently the layers predict diffusion in Table 2. (In Supplementary Appendix Figure S1 we also plot the 90% and 95% confidence intervals).

The advice layer stands out as most predictive, and we also see that the kerorice and social layers are also significantly predictive. Notably, in line with what we observed in the correlations and principal component, jati explains the least of the variation and is not significant.

Given how correlated the layers are, we also perform a LASSO (ℓ_1-penalized) regression to select a sparse set of relevant variables that explain the diffusion. We then use post-LASSO least squares to estimate the effect on diffusion of seed set centrality under the selected layer(s).

TABLE 2. Seed Set Diffusion Centrality

	No. Calls Received							
	1	2	3	4	5	6	7	8
Union	1.110							
	(1.752)							
	[0.529]							
Intersection		2.220						
		(2.200)						
		[0.317]						
Advice			6.410					
			(2.416)					
			[0.010]					
Kero/Rice				5.466				
				(2.326)				
				[0.022]				
Social					4.266			
					(1.820)			
					[0.022]			
Backbone						2.836		
						(2.203)		
						[0.203]		
Jati							1.161	
							(1.559)	
							[0.459]	
Decision								3.137
								(2.226)
								[0.164]
Num.Obs.	68	68	68	68	68	68	68	68
R2	0.110	0.139	0.313	0.254	0.194	0.147	0.110	0.161
Dep Var mean	8.691	8.691	8.691	8.691	8.691	8.691	8.691	8.691

Note: Robust standard erros are given in paranthesis and p-values in square brackets. Controls added for number of households, its powers, and dummy for number of seeds in the village. Exogenous variables are the sum of Diffusion Centrality for seeds in each village for the layer. Exogenous variables have been standardized.

The regression of interest is given by

$$y_v = \alpha + \sum_{\ell} \beta^{\ell} \cdot DC_v^{\ell} + X_v \Gamma^{\ell} + \epsilon_{v,\ell} \tag{2.2}$$

where the variables are as described in (2.1) and the distinction from (2.1) where we ran a separate regression for each layer is that this regression contains all layer variables.

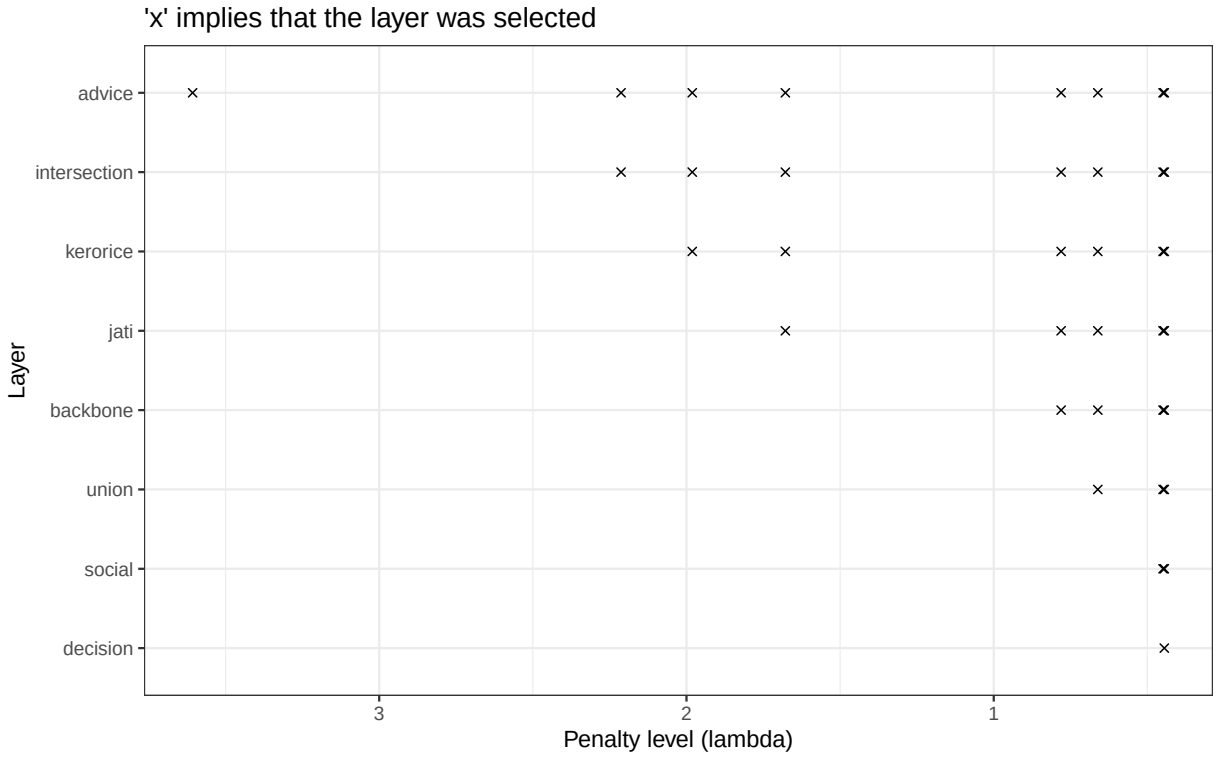


FIGURE 4. Lasso Selection of Layers in Predicting Diffusion

We are interested in which of the β^ℓ are estimated to be non-zero and the consistent estimates of these parameters.

A complication is that in order to be consistent LASSO requires a condition called irrepresentability, which is essentially that regressors of interest are not excessively correlated (Zhao and Yu, 2006). In our setting, this requirement fails. An approach to overcome this problem is to use the Puffer transformation developed by Rohe (2015) and Jia et al. (2015) which mechanically recovers irrepresentability when the number of observations exceeds the number of variables. Though the regressors $(DC_v^\ell)_{v,\ell}$ have correlated columns, by appropriately pre-conditioning the data matrix, we can force the columns of the data matrix to be orthogonal and therefore irrepresentable. Puffer-LASSO then recovers the set of relevant variables with probability tending to one exponentially fast in the number of observations, with consistent parameter estimates, that are with probability approaching one asymptotically normally distributed (Javanmard and Montanari, 2013; Jia et al., 2015; Taylor and Tibshirani, 2015; Lee et al., 2016; Banerjee et al., 2024a).

We see the results in Figure 4 where we plot which layer is selected by the LASSO as we increase the penalty level.

TABLE 3. Post Puffer Lasso OLS: Seed Set DC

	No. Calls Received
	1
Advice	5.564
	(2.117)
	[0.011]
Num.Obs.	68
R2	0.233
Dep Var mean	8.691

Robust Std.Errors are given in parenthesis, while p-values are given in square brackets. All exogenous variables have been standardized. Following set of variables were included in the lasso: Union, Intersection, Kerorice, Social, Decision, Advice, Jati and Backbone. The lasso at penalty level $\lambda = 2.95$ selected *advice*.

TABLE 4. F-test for the layers

layer	df	R.sq.	F-stat	p-val	F-stat marginal	p-val marginal
advice	1	0.233	20.057	0.000		
intersection	2	0.276	3.888	0.053	3.888	0.053
kerorice	3	0.281	2.134	0.127	0.415	0.522
jati	4	0.325	2.844	0.045	4.059	0.048
backbone	5	0.336	2.415	0.058	1.112	0.296
union	6	0.340	1.971	0.096	0.304	0.583
social	7	0.342	1.657	0.147	0.214	0.645
decision	8	0.344	1.419	0.215	0.138	0.712

We find that at the highest penalty level, the only network layer selected is advice, with the post puffer LASSO OLS regression in Table 3 showing a 64% increase in diffusion relative to the mean ($p = 0.011$). Despite the fact that multiple layers are useful in explaining diffusion, the backbone was not the most useful, nor was the union or intersection.

The fact that centrality in the advice layer is singled out as the best predictor of diffusion once there is a sufficiently high penalty does not mean that the other layers are not impacting diffusion. In fact, a combination of the layers still provides significantly more prediction than just the advice layer, as we see in Table 4.

Table 4 presents both cumulative and marginal F-tests as variables are added in the order selected by Lasso. Here we see, for instance, that adding intersection is marginally significant above advice, and then adding kerorice and jati yields a more complete model,

with an improvement significant at the 5 percent level.[12] Thus, even though jati is a poor substitute for other layers, it is still a useful complemment to them in predicting diffusion.

2.4. **How the Level of Multiplexing Affects Diffusion.**

Next, we examine how diffusion depends on the extent to which the layers in a village are multiplexed. That is, do villages that have greater correlation in their layers experience higher or lower levels of diffusion? To do this, we first have to define a measure of the extent to which a village is multiplexed.

To proceed, first we define a multiplexing score for household i in a village v as

$$m_{i,v} := \frac{\sum_j \left(\sum_\ell g^\ell_{ij,v}/L\right)}{\sum_j \mathbf{1}\{\sum_\ell g^\ell_{ij,v} > 0\}}.$$

The multiplexing score for a household i measures what fraction of distinct types of relationships it has when averaged across neighbors. The numerator is ratio of total number of links household i has across all relationship types to each of its unique neighbors divided by L and summed across neighbors. The denominator counts i's neighbors. For example, $m_{i,v} = 1$ if whenever household i has a relationship with some other household j, then it has all possible relationships with that other household. If, in contrast there is no multiplexing, then this measure would be $1/L$.

We then aggregate to the village level by taking $m_v := \frac{1}{n_v}\sum_i m_{i,v}$. Further, we define a dummy variable for having more than the median amount of multiplexing in the sample,

$$\text{High Mpx}_v := \mathbf{1}\left\{m_v > \text{median}(m_{1:v})\right\}.$$

Our regression of interest is

$$y_v = \alpha + \beta \cdot DC_v^{advice} \times \text{High Mpx}_v + \tau \cdot DC_v^{advice} + \eta \cdot \text{High Mpx}_v + X_v\Gamma + \epsilon_v. \tag{2.3}$$

where DC_v^{advice} denotes the diffusion centrality of seed set in village v for the "advice" layer (given that it was singled out as the best predictor of diffusion).

As before, τ captures the return to increasing the diffusion centrality of the seed set. Since information is seeded in all networks, η captures how the extent of diffusion changes with the worst possible seeding (the theoretical intercept). The coefficient β captures how incrementally improving seeding differentially affects the extent of diffusion as a function of the multiplexing.

[12]In Appendix Table S3 we exclude the extra layers of intersection, union, and backbone, which are "constructed" layers that are derived from these basic layers. F-tests include the basic layers in the order selected by the Lasso.

We note that the interaction term $DC_v^{advice} \times \text{High Mpx}_v$ is important, and its coefficient of primary interest, since in villages with low seed set centrality there is very little diffusion, and hence multiplexing cannot really make any difference. Thus, multiplexing's marginal impact (positive or negative) should be most pronounced in settings where the seed set centrality is high.

Indeed, this is confirmed in Table 5.

TABLE 5. Multiplexing and Diffusion

	Calls per Household
	(1)
High Multiplexing	−0.023
	(0.016)
	[0.164]
Seed Set Centrality	0.052
	(0.016)
	[0.002]
High Multiplexing X Seed Set Centrality	−0.039
	(0.017)
	[0.022]
Num.Obs.	68

Robust Std.Errors are given in parenthesis, while p-values are given in square brackets. Seed Set Centrality comes from the "advice" layer and has been standardized. Controls for number of seeds and average total degree across network layers have been added.

As one would expect, the coefficient on seed set centrality is positive and significant. We also find both $\beta < 0$ and $\eta < 0$. Qualitatively, $\eta < 0$ means that more multiplexed networks generate less diffusion, of course with the caveat that these villages could be different for other reasons, and this coefficient is not significant. Importantly, the coefficient $\beta < 0$ implies that settings with more central seeding—and hence higher levels of diffusion—end up having their diffusion impeded by multiplexing.

3. A Theory of Diffusion and Multiplexing

Having established that: (i) the network layers are distinct but strongly correlated/multiplexed, (ii) they are differently predictive of diffusion, (iii) multiple layers are predictive of diffusion, and (iv) more multiplexed villages experience less diffusion, we now develop a theory that helps us understand how and why multiplexing affects diffusion.

We approach the problem from two different levels. The first is the individual level. Here, we examine how an individual's probability of becoming infected depends on their multiplexing (for any given probability of infection among neighbors). The second aggregates this to the societal infection level. There we use the results about individuals as a key lemma in analyzing a canonical SIS contagion process.

We also consider both simple diffusion/contagion in which a single contact is enough for an individual to become infected, as well as complex diffusion in which multiple

contacts are needed. We take these in turn, each time beginning with a result about individual infection probabilities and then aggregating to the societal level.

We begin by outlining a general model of multiplexed diffusion.

3.1. **A Model of Diffusion with Multiplexing.**

We study diffusion/contagion in a society consisting of a finite set of individuals $N = \{1, \ldots, n\}$. Each individual has relationships captured via layers $\{1, \ldots, L\}$, with a generic layer represented by ℓ. In each layer ℓ, the interactions between individuals are described by a (possibly directed) network with adjacency matrix $g^\ell \in \{0,1\}^{n \times n}$, such that $g^\ell_{ij} = 1$ if there is a link from i to j in layer ℓ (e.g., i follows j or could be infected by j, etc.), and 0 otherwise. We define the multigraph consisting of L layers to be $g = \langle g^1, g^2, \ldots, g^L \rangle$.

Let $\mathcal{L}_{ij} = \{\ell | g^\ell_{ij} = 1\}$ denote the set of layers in which there is a directed link from i to j. The set of all neighbors for a given node i is denoted $\mathcal{N}_i = \{j | \mathcal{L}_{ij} \neq \emptyset\}$.

To track infection across time, we index discrete periods via $t \in \{0, 1, 2, \ldots\}$. At each point in time an individual in the network is in one of two states: Susceptible (S) or Infected (I). The status of individual i at time t is denoted by the random variable $x_i(t)$. If $x_i(t) = 1$, individual i is infected at time t; if $x_i(t) = 0$, individual i is susceptible at time t. The state of the society at time t is given by the vector $x(t) = \langle x_1(t), x_2(t), \ldots, x_n(t) \rangle \in \{0,1\}^n$.

At each time t, an individual's state can change based on the infection status of its neighbors. A susceptible individual i becomes infected if it receives at least τ infection transmissions from its infected neighbors in a given time period. An infected individual recovers (and becomes susceptible again) randomly with a probability δ at the end of a period. If $\tau = 1$ then this is a standard (simple) contagion process, while with a threshold $\tau > 1$ then this involves a more complex contagion (Granovetter, 1978; Centola, 2010), and becomes a game on a network (Jackson and Zenou, 2014).

Our model of contagion on a multiplexed network involves two additional features.

First, we allow different layers to have different contact probabilities, which is needed given the heterogeneity in prediction we saw in Section 2.3. Specifically, for each layer ℓ, let $q_\ell \in (0,1)$ be the marginal probability of infection transmission from an infected individual to a susceptible one if they are connected via that layer.

Second, given that individuals could be connected via multiple layers, we need to define how transmission takes place through multiple layers. Let $x_{ij}(t)$ represent the number of infection transmissions at time t from an infected node j to a susceptible node i. We denote the probability of receiving infection transmissions across k layers

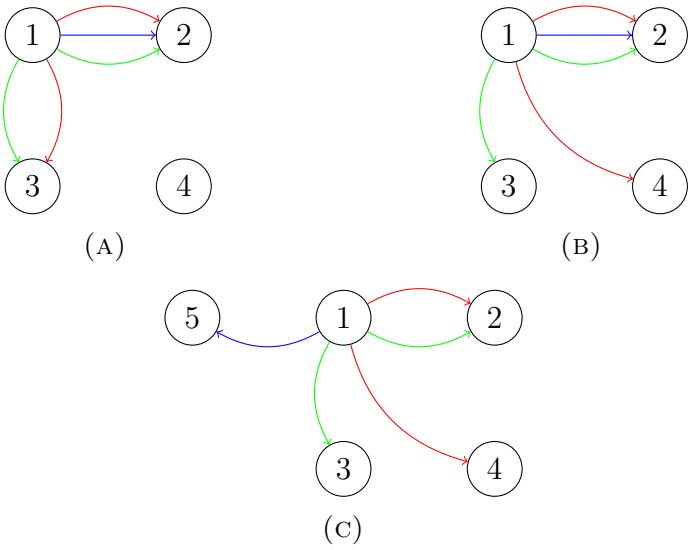


FIGURE 5. Node 1's relationships are successively less multiplexed moving from panel (A) to (C)

from node j, given $\mathcal{L}_{ij}$ by

$$f(k; \mathcal{L}_{ij}) := P(x_{ij} = k \mid \mathcal{L}_{ij}).$$

$f(k; \mathcal{L}_{ij})$ can capture arbitrary patterns of correlation in infection transmission through multiple layers. For example, if there is a positive correlation in transmission across layers then if two nodes are connected by layers $\{A, B\}$, $f(2, \{A, B\}) \geq q_A q_B$.

The probability that a susceptible individual i becomes infected at time t given the infection status of its neighbors at time $t-1$ is

$$P\left(\sum_{j \in \mathcal{N}_i} x_{ij} x_j(t-1) \geq \tau\right)$$

3.1.1. *Comparisons of Multiplexing.* Since it is not always possible to order two multigraphs in terms of multipexing, we define a partial order on the set of multigraphs. We begin with an example illustrating the concept in Figure 5.

In Figure 5 panel A we depict a multigraph with 5 nodes and 3 layers. In panel B, by moving node 1's link in layer *red* from node 3 to node 4, we arrive at a graph that is less multiplexed, while maintaining the same out-degree. Similarly, in panel C, we again move node 1's link in layer *blue* from node 2 to node 5, giving us a less multiplexed network as compared to panel B.

We begin by defining a *local multiplexity dominance relation*, denoted as $\prec$. For two multigraphs g and $\widehat{g}$, we say $\widehat{g} \prec g$—that is $\widehat{g}$ is locally less multiplexed than g—if $\widehat{g}$ can be obtained from g by removing a link in some layer ℓ between nodes i and j and adding a new link in that same layer to another neighbor k, where i's connections to k were a strict *subset* of those with j (excluding ℓ) to start with. This means that: (i) $\mathcal{L}_{ik}(g) \subsetneq \mathcal{L}_{ij}(g) \setminus \{\ell\}$, (ii) $g^{\ell}_{ik} = 0 = \widehat{g}^{\ell}_{ij}$ and $\widehat{g}^{\ell}_{ik} = 1 = g^{\ell}_{ij}$, and (iii) for all other links $\widehat{g}$ and g coincide.

Given that thelocal multiplexity dominance relation is acyclic (Appendix Proposition 5), we define the *less multiplexed relation*, denoted by $\overline{\prec}$ as the transitive closure of $\prec$. That is, we say that $\widehat{g}\overline{\prec}g$ if there exists a finite sequence of multigraphs $g_1, g_2, \ldots, g_k$ such that $\widehat{g} = g_1 \prec \cdots \prec g_k = g$. The relation $\overline{\prec}$ forms a partial order on the set of multigraphs.

We also define this with respect to a given node i, where $\widehat{g}\overline{\prec}_i g$ if $\widehat{g}\overline{\prec}g$ and $\widehat{g}_i \neq g_i$ (where g_i denotes the row(s) associated with node i).

3.2. **Multiplexing Impedes Simple Diffusion and Contagion.**

We first analyze the case of simple contagion, $\tau = 1$. We focus on the case of two layers as this captures all of the intuition.

3.2.1. *Infection of an Individual.*

To gain intuition for how multiplexing impedes diffusion, consider some node i that is connected on both layers (denoted by A, B) to node j in g but not to node k. Then the change from g to $\widehat{g}$ is to remove one of the layers of connection from j and add it to k. Suppose that both j and k are independently infected with probability ρ, and similarly for any of i's other connections.

The probability of i becoming infected is higher from two un-multiplexed links if and only if:

$$q_A\rho + q_B\rho - f(2, \{A, B\})\rho \leq q_A\rho + q_B\rho - q_Aq_B\rho^2,$$

where A, B are the layers and this restricts attention to probability of i getting infected by one of these two nodes (presuming no infection from any of the other nodes, which is the only case that matters and for which the conditional probability has not changed). This simplifies to

$$q_Aq_B\rho \leq f(2, \{A, B\}). \tag{3.1}$$

A sufficient condition for this to hold is that $q_Aq_B \leq f(2, \{A, B\})$, or that transmission is independent across layers. For $\rho < 1$, there can even be negative correlation, as long as transmission across layers is not too negatively correlated. Intuitively, strong

negative correlation across layers provides a large advantage to multiplexing in the case of simple diffusion since it minimizes the chance of having two transmissions from the same neighbor and spreads those probabilities out. As long as such strong negative correlation does not hold, multiplexing hurts diffusion. This is because multipexling reduces the diversification of contacting multiple individuals, which increases the chances of hitting one that is infected.

Proposition 1. *Under simple contagion ($\tau = 1$), if $\widehat{g} \overline{\prec}_i g$, and each of i's neighbors are infected independently with probability $\rho > 0$, and i is susceptible, then i is more likely to be infected under the less multiplexed network g than under $\widehat{g}$* if and only if *transmission is not too negatively correlated across layers (condition 3.1), with the reverse holding if condition 3.1 fails.*

3.2.2. *Multiplexing and Overall Infection in the SIS Model.*

Proposition 1 suggests that the overall infection rate in a variety of contagion processes will be higher on less multiplexed networks. This is true, and we state it for the case of the SIS model.

To perform this analysis, we extend the mean-field techniques that are standardly used to solve the SIS model in one dimension (e.g., see Pastor-Satorras and Vespignani (2000); Jackson (2008)), to study it under multiplexing.

A given node i's connections are described by a vector of subsets of layers, $D_i = (D_{i1}, \ldots, D_{iK})$, where $K \leq n-1$ is the total number of neighbors of the node, and $D_{ik} \subset \{1, \ldots, L\}$ is the set of layers that i is connected to its kth neighbor on, where each $D_{ik} \neq \emptyset$.

A sufficient statistic for D_i for the mean-field analysis is a triple $\hat{D}_i = (\hat{D}_{iA}, \hat{D}_{iB}, \hat{D}_{i,AB})$, which is the number of connections that i has that are just on layer A, just on layer B, and on both layers, respectively.

The distribution of $\hat{D}$ across the population is described by the function $P(\hat{D})$, which we assume to have finite support. The steady-state infection rate of nodes with connection profile $\hat{D}$ is denoted $\rho(\hat{D})$. The population infection rate is then defined by

$$\rho = \sum_{\hat{D}} P(\hat{D})\rho(\hat{D}). \tag{3.2}$$

The probability that a susceptible node with connections $\hat{D} = (\hat{D}_A, \hat{D}_B, \hat{D}_{AB}) \neq (0,0,0)$ becomes infected, in steady state, is then

$$1 - (1-\rho q_A)^{\hat{D}_A}(1-\rho q_B)^{\hat{D}_B}(1-\rho[(q_A + q_B - f(2,\{A,B\})])^{\hat{D}_{AB}}. \tag{3.3}$$

In the mean-field analysis the steady state equation for nodes with connections $\hat{D} = (\hat{D}_A, \hat{D}_B, \hat{D}_{AB}) \neq (0,0,0)$, as a function of the overall infection rate ρ, is the solution to

$$\rho(\hat{D})\delta = \tag{3.4}$$
$$(1-\rho(\hat{D}))\left[1-(1-\rho q_A)^{\hat{D}_A}(1-\rho q_B)^{\hat{D}_B}(1-\rho[(q_A+q_B-f(2,\{A,B\})])^{\hat{D}_{AB}}\right].$$

A steady-state is then a joint solution to (3.2) and (3.4) for each $\hat{D}$ in the support of P. 0 is always a solution, and for some P there also exists a positive solution.

We extend the partial order we defined in 3.1.1 as follows. We say that a distribution P' is less multiplexed that P, denoted by $P' \overline{\prec} P$, if there exists $\hat{D}$ and $\hat{D}'$ such that

- $\hat{D}'_A = \hat{D}_A + 1$,
- $\hat{D}'_B = \hat{D}_B + 1$,
- $\hat{D}'_{AB} = \hat{D}_{AB} - 1$,
- $P'(\hat{D}') + P'(\hat{D}) = P(\hat{D}') + P(\hat{D})$, and
- $P'(\hat{D}') > P(\hat{D}')$.

We define $\overline{\prec}$ as the transitive closure of $\prec$.

Proposition 2. *Consider a simple contagion process ($\tau = 1$) process with transmission probabilities described by $(q^\ell)_\ell \in (0,1)^L$ and f, and a recovery rate $\delta \in (0,1)$ and two distributions of connections P' and P that each have a positive steady-state infection rate. If $P' \overline{\prec} P$, then the positive steady-state infection rate under P' is higher than that under P* if and only if *transmission is not too negatively correlated (condition 3.1) at the positive infection rate of P.*

Proposition 2 implies that multiplexing has consequences, which are good or bad depending on the context of whether diffusion is socially desired (e.g., information about a beneficial program) or not (e.g., spread of a disease). Given the many things that may lead to multiplexing, this implies that the things that cause people to layer their networks have strong implications for diffusion processes. This also means that networks whose layers are optimized for one reason can be suboptimal for another.

3.3. **Multiplexing and Complex Diffusion.**

The results on simple contagion are unambiguous: multiplexing impedes simple diffusion/contagion except in extreme cases of negatively correlated transmission probabilities. It turns out that complex contagion is more complex. Multiplexing can both enhance and impede diffusion, depending on the circumstances.

In the case of complex diffusion there are now competing forces of multiplexing. One force is the same force driving how less multiplexing impacts the simple contagion process—diversifying links leads to a higher probability of at least some links hitting

infected individuals. However, now the counter-force is that conditional upon hitting an infected individual, multiplexing leads to higher probabilities of multiple transmissions than spreading those links to other individuals who might be uninfected.

Again, to keep the analysis as uncluttered as possible, we focus on the case of two layers.

Finally, we also consider a case where the correlation in transmission across layers is neither too high nor too low, so that there is an ε, to be determined in the proofs of the propositions below, for which $(1+\varepsilon)q^Aq^B \geq f(2,\{A,B\}) \geq q^Aq^B$. Of course, a sufficient condition for this to hold is independent transmission. This condition is needed as with excessive positive or negative correlation in transmission, strange patterns of transmission as a function of multiplexing can occur. For instance, if transmission is perfectly positively correlated, then one is always more likely to get two transmissions from a single multiplexed connection than two unmultiplexed connections, but is always more likely to get one transmission from the reverse. This then implies that the optimal configuration of connections depends on whether τ is even or odd, in complicated ways as a function of a node's overall degrees in each layer. While this integer detail is interesting, the following results highlight the more fundamental forces of multiplexing.

Proposition 3. *Consider a complex contagion ($\tau > 1$), a susceptible node i such that i's neighbors are infected independently with probability $\rho > 0$, and two networks such that $\widehat{g} \overline{\prec}_i g$. Also suppose that $\sum_{\ell,j} g_{ij}^{\ell} > \tau$, so that i has more than enough connections to become infected.*

There exist $0 < \underline{\rho} < \overline{\rho} < 1$ such that

- *if $\rho, q^A, q^B > \overline{\rho}$, then i is* less *likely to be infected under the more multiplexed network g than under $\widehat{g}$, and*
- *if $\rho, q^A, q^B < \underline{\rho}$, then i is* more *likely to be infected under the more multiplexed network g than under $\widehat{g}$.*

The intuition behind the result is as follows. There exist i, j, k such that under g, node i is connected to j on two layers and k on none, while under $\widehat{g}$, node i is connected to j on one layers and k on the other. The cases in which this difference can be pivotal are when the other connections to other nodes have led to either $\tau-1$ or $\tau-2$ transmissions. With a very high infection and transmission rate among neighbors, the $\tau-1$ case is more likely and then this looks more like simple contagion, and so less multiplexing leads to a higher chance of infection. Under a very low infection rate among neighbors, the $\tau-2$ case is more likely and one needs two infections, and that is very unlikely across two neighbors, but much more likely to happen with just one neighbor. Thus, more multiplexing leads to a higher chance of infection.

Interestingly, as we will see in simulations below, in the intermediate range these forces can interact non-monotonically, which is why we have a gap between the upper and lower bounds.

We now state how this translates into an aggregate infection rate.

Proposition 4. *Consider a complex contagion ($\tau > 1$), with nonnegative correlation in transmission across layers, so that $f(2, \{A, B\}) \geq q^A, a^B$, and two distributions such that $P' \trianglerighteq P$ and both have positive steady-state infection rates. Also suppose that $\hat{D}_A + \hat{D}_B + 2\hat{D}_{AB} > \tau$ for each $\hat{D}$ in the distribution P, so that each node has more than enough connections to become infected. There exist $0 < \underline{\rho} < \overline{\rho} < 1$ such that*

- *if $q^A, q^B < \underline{\rho}$ and δ is sufficiently high, then the steady-state infection is higher for P than P', and*
- *if $q^A, q^B > \overline{\rho}$ and δ is sufficiently low, then the steady-state infection is lower for P than P'.*

Note that the steady-state infection of every connection type shares the same ordering as the overall infection rate.

3.4. **Simulations.** The theoretical results are based on random networks that in the large end up having well-defined statistical properties. To see how the results work in smaller empirical networks, we run simulations on the networks from the RCT villages. We simulate a Susceptible-Infected-Susceptible (SIS) diffusion process for the cases of both simple and complex diffusion. In order to compare across similar sized networks where only multiplexing is changing, we begin from a given village network and compare how diffusion varies when we use different pairs of networks (which end up empirically having different multiplexing rates) to make a clean multiplexing comparison. We then simulate diffusion across multiple villages and iterations of the simulations.

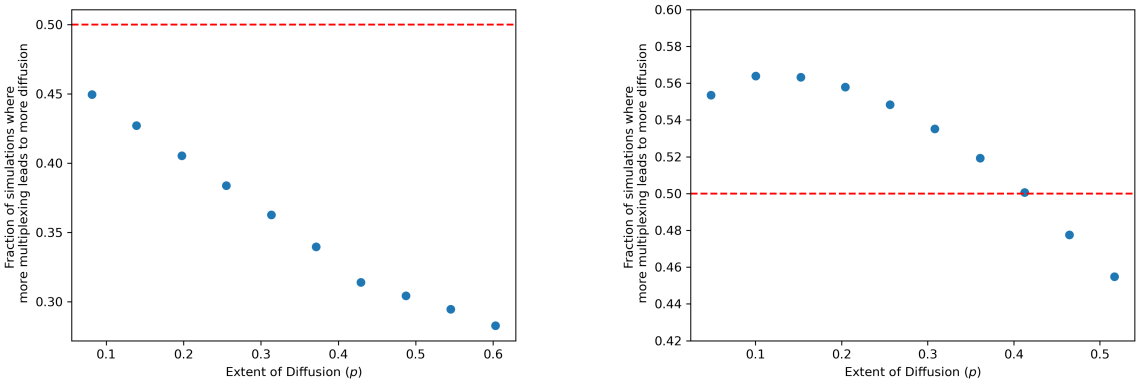
In particular, for each village, we begin by picking three empirical adjacency matrices representing different network layers sorted in decreasing order of their average out-degree: A_1, A_2, and A_3. We then pair A_1 with A_2 for one simulated diffusion, and A_1 with A_3 for the other. To ensure that the average out-degree is comparable across the networks, we prune at random the links in A_2 to match the average out-degree of A_3, resulting in a pruned network $A_2^{'}$. We construct two multiplexed networks: g', by combining A_1 and $A_2^{'}$, and g, by combining A_1 with the A_3. The process is presented in full detail in the Appendix (Algorithm 2).

The diffusion process (also presented in full detail in the appendix in Algorithm 1), is as follows. First, a susceptible node can get message transmissions from each infected neighbor in each layer, i.i.d., with probability q in each period. Second, a susceptible

node gets infected only if it receives at least $\tau \geq 1$ contacts in a given time period and the count resets in each time period. Third, in each period an infected node transitions back to being susceptible with probability δ. If the share of infected nodes does not change beyond a small thresh-old between two consecutive iterations, we terminate the simulation. In our simulations we use $\tau = 1$ for simple diffusion and $\tau = 2$ for complex. In each simulation we set the number of randomly selected seeds in the initial period to be $\lfloor \sqrt{n} \rfloor$, where n is the number of households in the network. For both, simple diffusion ($\tau = 1$) as well as complex diffusion ($\tau = 2$) we run simulations on a grid of $(q, \delta) \in [0.1, 0.5] \times [0.1, 0.5]$. We run the diffusion simulations 500 times for each village across both multiplexed networks g, g' described above. We report the averages across all 70 villages.

Given that these are smaller networks, some simulations end up randomly having more or less diffusion in any given run across the two comparison networks. Thus, we tabulate the fraction of simulation runs for which more multiplexing is associated with more diffusion.

The results are presented in Figure 6



(A) Simple Contagion ($\tau = 1$) (B) Complex Contagion ($\tau = 2$)

FIGURE 6. Diffusion Simulations

In Figure 6 we plot the fraction of simulation runs where more multiplexing leads to more diffusion against the extent of diffusion in the network p. In panel A, we plot the results for simple (pure) diffusion. We find that higher multiplexing is consistently associated with lower diffusion levels, as in our theoretical results.

In panel B we see the non-monotonicity from the countervailing forces in complex diffusion that we mentioned in Section 3.3. We also see a confirmation of the theoretical results. At low levels of diffusion, the steady state diffusion is increasing in multiplexing and for high diffusion levels, the steady state diffusion is decreasing in multiplexing.

4. Concluding Remarks

Despite the many case studies of multiplexed networks, there has been little work on how it affects behavior and outcomes. We have shown that multiplexing systematically impacts diffusion, both in theory and in an RCT.

Our findings highlight the need for future work on incentives to multiplex and resulting consequences. There are several immediate directions to explore. For example, our results suggest that the deeper the need to form reinforced or supported relationships, the greater the inefficiencies in certain domains. So those who are under weaker institutions or have limited resources may face a greater need to multiplex relative to their richer counterparts. As a result, they may have worse access to information on the one hand, and more be more susceptible to the spread of social norms.

There is also a need for further development of measures and methods of analyzing multiplexed networks. We defined one of many potential measures of how multiplexed a network is, as well as one of many potential partial orders. Understanding which measures are most appropriate in which settings, is a subject for further research.

To close, we show two other patterns that we found in the data. For both of these calculations we use the multiplexing score that we defined in Section 2.4:

$$m_{i,v} := \frac{\sum_j \left(\sum_\ell g^\ell_{ij,v}/L\right)}{\sum_j \mathbf{1}\{\sum_\ell g^\ell_{ij,v} > 0\}},$$

where here i is either an individual or a household, depending on the analysis.

The first pattern is that higher degree households are less multiplexed. We restrict our attention to the RCT villages and specifically to the social and advice layers.

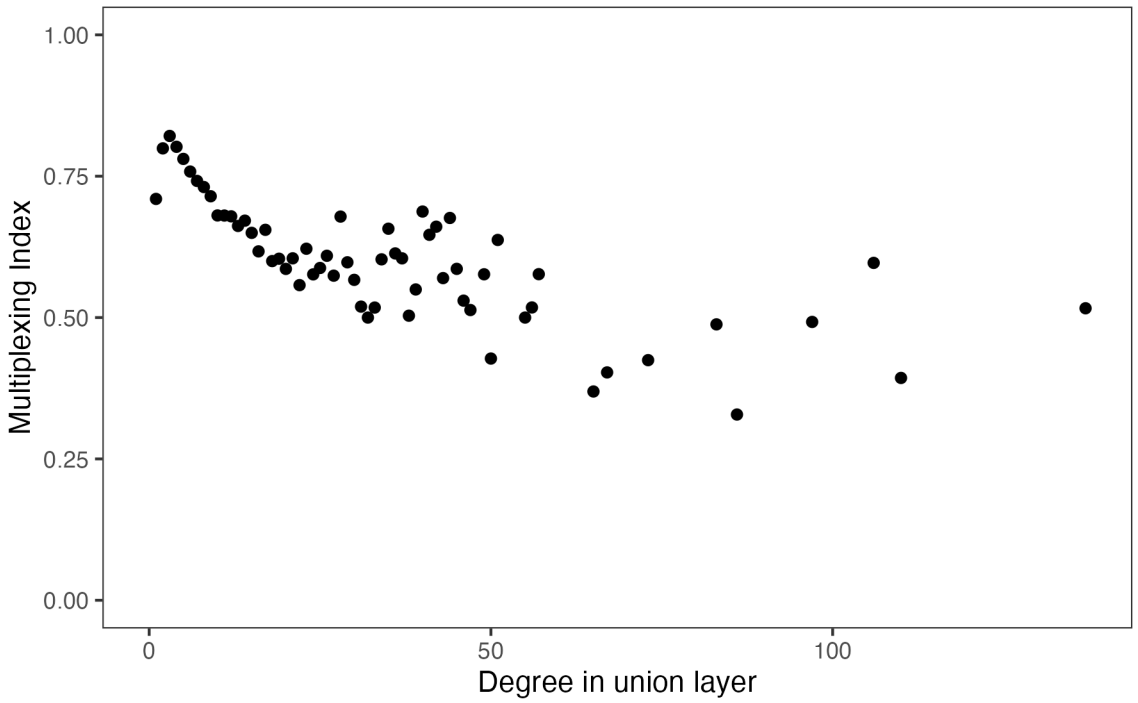


FIGURE 7. Multiplexing as a function of degree.

Figure 7 depicts a binned scatter plot where we can see that households that have higher degree (aggregated across layers) have lower levels of multiplexing.

The second pattern is that females' networks are significantly more multiplexed than males'.[13] Here we use the individual network data from Microfinance villages collected as part of the Wave I of data collection in Banerjee et al. (2013).[14] We use the social, kerorice, advice, decision, money, temple, and medic layers. For each village v, we aggregate this score at the gender level: $m_{a,v} = \frac{1}{n_a} \sum_{i \in I_a} m_{i,v}$, where $a \in \{male, female\}$. Figure 8 exhibits the density curves for these multiplexing scores across the villages. The curve for women is effectively a right-shift, showing that women's networks are systematically more multiplexed.

[13]See Beaman and Dillon (2018) for an exploration of heterogeneity in diffusion by gender, though the focus is not on multiplexing.

[14]We have individual-level gender distinguished data in the Wave I network survey, which elicited links from 46% of the households, giving us information on 70.84% of the links.

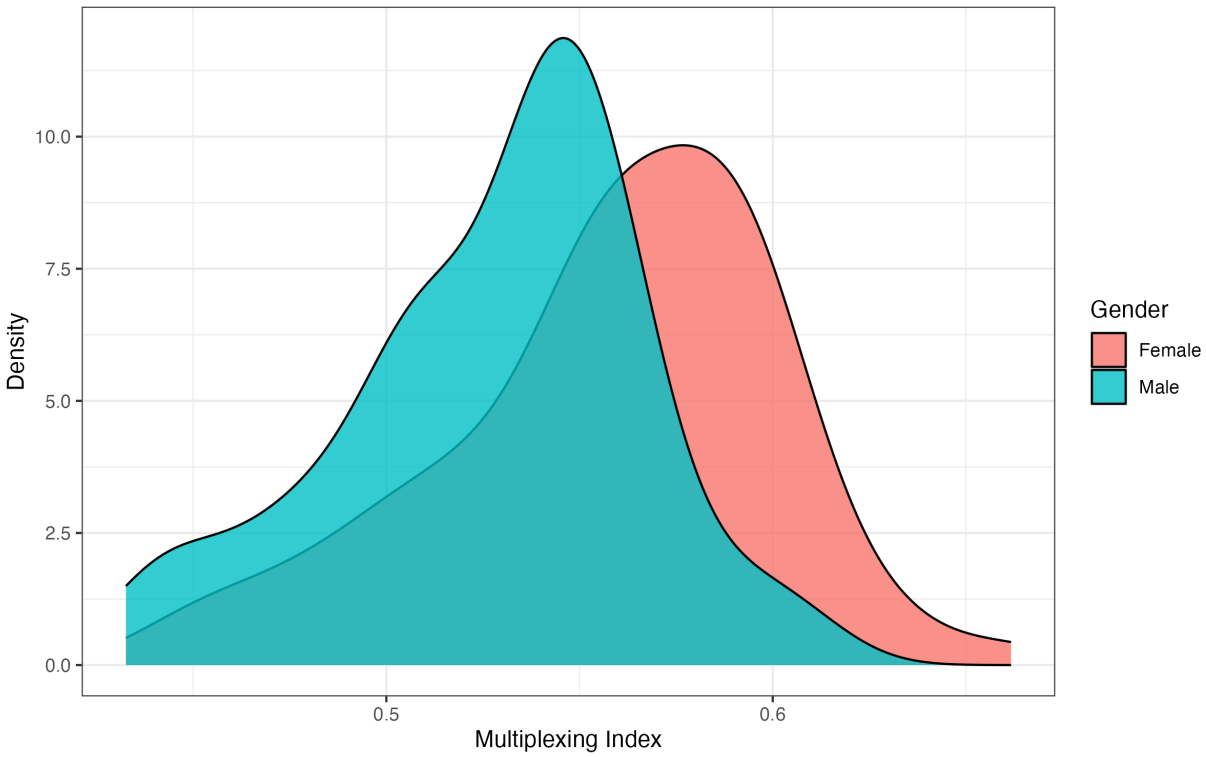


FIGURE 8. Multiplexing by gender.

Note that women in rural Indian communities often marry across village boundaries (though frequently still within the constraints of caste/jati endogamy) and most of these marriages are virilocal—requiring the wife to move into husband's house (Rosenzweig and Stark, 1989; Rao and Finnoff, 2015). As a consequence, they often rely on affinal kin and over time need to "rebuild" their networks (Hruschka et al., 2023). This is in conjunction with the expectation that these women take on various responsibilities, including agricultural work, managing the household, preparing meals, and raising children. Such constraints on available relationships while serving multiple roles can plausibly result in high levels of multiplexing, an interesting subject for further research.

REFERENCES

AMBRUS, A., M. MOBIUS, AND A. SZEIDL (2014): "Consumption risk-sharing in social networks," *American Economic Review*, 104, 149–182. 1

ATKISSON, C., P. J. GÓRSKI, M. O. JACKSON, J. A. HOŁYST, AND R. M. D'SOUZA (2019): "Why understanding multiplex social network structuring processes will help us better understand the evolution of human behavior," *forthcoming: Evolutionary Anthropology*. 1

Banerjee, A., A. G. Chandrasekhar, S. Dalpath, E. Duflo, J. Floretta, M. O. Jackson, H. Kannan, F. N. Loza, A. Sankar, A. Schrimpf, et al. (2024a): "Selecting the Most Effective Nudge: Evidence from a Large-Scale Experiment on Immunization," *Econometrica, forthcoming*. 2.3

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2019): "Using gossips to spread information: Theory and evidence from two randomized controlled trials," *The Review of Economic Studies*, 86, 2453–2490. 1, 2.1.2, 2.3

Banerjee, A. V., E. Breza, A. G. Chandrasekhar, E. Duflo, M. O. Jackson, and C. Kinnan (2024b): "Changes in Social Network Structure in Response to Exposure to Formal Credit Markets," *Review of Economic Studies*, 91:3, 1331–72. 1, 2.1.1

Banerjee, A. V., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013): "Diffusion of Microfinance," *Science*, 341, DOI: 10.1126/science.1236498, July 26 2013. 1, 2.1.1, 2.3, 4

Beaman, L. and A. Dillon (2018): "Diffusion of agricultural information within social networks: Evidence on gender inequalities from Mali," *Journal of Development Economics*, 133, 147–161. 13

Bianconi, G. (2018): *Multilayer networks: structure and function*, Oxford university press. 1

Billand, P., C. Bravard, S. Joshi, A. S. Mahmud, and S. Sarangi (2023): "A model of the formation of multilayer networks," *Journal of Economic Theory*, 213, 105718. 1

Boccaletti, S., G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin (2014): "The structure and dynamics of multilayer networks," *Physics reports*, 544, 1–122. 1

Centola, D. (2010): "The Spread of Behavior in an Online Social Network Experiment," *Science*, 329: 5996, 1194–1197, DOI: 10.1126/science.1185231. 3.1

Cheng, C., W. Huang, and Y. Xing (2021): "A theory of multiplexity: Sustaining cooperation with multiple relations," *Available at SSRN 3811181*. 1

Contractor, N., P. Monge, and P. M. Leonardi (2011): "Network Theory—multidimensional networks and the dynamics of sociomateriality: bringing technology inside the network," *International Journal of Communication*, 5, 39. 1

Dickison, M. E., M. Magnani, and L. Rossi (2016): *Multilayer social networks*, Cambridge University Press. 1

Fafchamps, M. and F. Gubert (2007): "Risk sharing and network formation," *American Economic Review*, 97, 75–79. 1, 2.1.1

Granovetter, M. S. (1978): "Threshold models of collective behavior," *American journal of sociology*, 83, 1420–1443. 3.1

Hruschka, D. J., S. Munira, and K. Jesmin (2023): "Starting from scratch in a patrilocal society: how women build networks after marriage in rural Bangladesh," *Philosophical Transactions of the Royal Society B*, 378, 20210432. 4

Hu, Y., D. Zhou, R. Zhang, Z. Han, C. Rozenblat, and S. Havlin (2013): "Percolation of interdependent networks with intersimilarity," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 88, 052805. 1

Jackson, M. O. (2008): *Social and economic networks*, Princeton: Princeton University Press. 3.2.2

Jackson, M. O., S. M. Nei, E. Snowberg, and L. Yariv (2024): "The Dynamics of Networks and Homophily," *SSRN Working Paper https://papers.ssrn.com/abstract=4256435*, . 1

Jackson, M. O. and Y. Zenou (2014): "Games on Networks," *Handbook of Game Theory, Elsevier, edited by Young, H.P. and Zamir, S.* 3.1

Javanmard, A. and A. Montanari (2013): "Model selection for high-dimensional regression under the generalized irrepresentability condition," in *Proceedings of the 26th International Conference on Neural Information Processing Systems-Volume 2*, 3012–3020. 2.3

Jia, J., K. Rohe, et al. (2015): "Preconditioning the lasso for sign consistency," *Electronic Journal of Statistics*, 9, 1150–1172. 2.3

Kivela, M., A. Arenas, J. P. Gleeson, Y. Moreno, and M. A. Porter (2014): "Multilayer Networks," *arXiv:1309.7233v4 [physics.soc-ph]*. 1

Kobayashi, T. and T. Onaga (2023): "Dynamics of diffusion on monoplex and multiplex networks: A message-passing approach," *Economic Theory*, 76, 251–287. 1

Larson, J. M. and P. L. Rodriguez (2023): "The risk of aggregating networks when diffusion is tie-specific," *Applied Network Science*, 8, 21. 1

Lee, J. D., D. L. Sun, Y. Sun, and J. E. Taylor (2016): "Exact post-selection inference, with application to the lasso," *The Annals of Statistics*, 44, 907–927. 2.3

Luo, W. and B. Li (2016): "Combining eigenvalues and variation of eigenvectors for order determination," *Biometrika*, 875–887. 2.2.3, 11

Morelli, S. A., D. C. Ong, R. Makati, M. O. Jackson, and J. Zaki (2017): "Empathy and well-being correlate with centrality in different social networks," *Proceedings of the National Academy of Sciences*, 114, 9843–9847. 1

Munshi, K. and M. Rosenzweig (2009): "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth," *mimeo*. 2.1.1

Pastor-Satorras, R. and A. Vespignani (2000): "Epidemic Spreading in Scale-Free Networks," *Physical Review Letters*, 86 (14), 3200–3203. 3.2.2

Rao, S. and K. Finnoff (2015): "Marriage migration and inequality in India, 1983–2008," *Population and Development Review*, 41, 485–505. 4

Rohe, K. (2015): "Preconditioning for classical relationships: a note relating ridge regression and OLS p-values to preconditioned sparse penalized regression," *Stat*, 4, 157–166. 2.3

Rosenzweig, M. R. and O. Stark (1989): "Consumption smoothing, migration, and marriage: Evidence from rural India," *Journal of political Economy*, 97, 905–926. 4

Sacerdote, B. (2001): "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly journal of economics*, 116, 681–704. 2.1.1

San Román, D. (2024): "Multiplexed Network Formation and Bonacich Centrality," *Available at SSRN 4704738*. 1

Simmel, G. (1908): *Sociology: Investigations on the Forms of Sociation*, Leipzig: Duncker and Humblot. 1

Taylor, J. and R. J. Tibshirani (2015): "Statistical learning and selective inference," *Proceedings of the National Academy of Sciences*, 112, 7629–7634. 2.3

Townsend, R. M. (1994): "Risk and Insurance in Village India," *Econometrica*, 62, 539–591. 1

Walsh, A. M. (2019): "Games on multi-layer networks," . 1

Wasserman, S. and K. Faust (1994): *Social Network Analysis*, Cambridge: Cambridge University Press. 1

Yağan, O. and V. Gligor (2012): "Analysis of complex contagions in random multiplex networks," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 86, 036103. 1

Zenou, Y. and J. Zhou (2024): "Games on Multiplex Networks," *Available at SSRN 4772575*. 1

Zhao, P. and B. Yu (2006): "On model selection consistency of Lasso," *The Journal of Machine Learning Research*, 7, 2541–2563. 2.3

Zhu, S.-S., X.-Z. Zhu, J.-Q. Wang, Z.-P. Zhang, and W. Wang (2019): "Social contagions on multiplex networks with heterogeneous population," *Physica A: Statistical Mechanics and its Applications*, 516, 105–113. 1

Appendix A. Appendix

A.1. **Proofs.**

Proof of Proposition 1: We adopt the notation from Proposition 2, as given independent probabilities of infection of neighbors, the probability that an individual with connection profile $\hat{D} = (\hat{D}_A, \hat{D}_B, \hat{D}_{AB})$ on network g becomes infected is then (from (3.3) given by

$$1 - (1 - \rho q_A)^{\hat{D}_A}(1 - \rho q_B)^{\hat{D}_B}(1 - \rho[(q_A + q_B - f(2, \{A, B\})])^{\hat{D}_{AB}}.$$

If the change is to network $\hat{g}$ in which this individual is less multiplexed then their connection profile is $(\hat{D}_A + a, \hat{D}_B + a, \hat{D}_{AB} - a)$ for some integer $a > 0$, and then their probability of being infected is

$$1 - (1 - \rho q_A)^{\hat{D}_A + a}(1 - \rho q_B)^{\hat{D}_B + a}(1 - \rho[(q_A + q_B - f(2, \{A, B\})])^{\hat{D}_{AB} - a}.$$

The second probability is larger than the first if and only if

$$(1 - \rho q_A)^a(1 - \rho q_B)^a(1 - \rho[(q_A + q_B - f(2, \{A, B\})])^{-a} < 1,$$

which simplifies to

$$(1 - \rho q_A)(1 - \rho q_B) < (1 - \rho[(q_A + q_B - f(2, \{A, B\})]).$$

This holds if and only if

$$\rho q_A q_B < f(2, \{A, B\}),$$

which is the claimed condition. ∎

Proof of Proposition 2: Following the argument from the proof of Proof of Proposition 1, for any ρ equation 3.4 has a higher solution for the less multiplexed type. Thus, starting with the steady state ρ for the more multiplexed distribution, the new rates for all individuals are weakly and sometimes strictly higher for the less multiplexed distribution. This leads to a higher ρ'. Iterating, this converges upward for all types to a limit which is the steady state. Conversely, if condition 3.1 is reversed, the convergence is downward for all types. ∎

Proof of Proposition 3: It is enough to consider an individual i with one change in their links where a multiplexed link to j is then split to two neighbors j, k, each of which is connected to i on a different layer and where i was initially not connected to k. Our focus is on the pivotal cases:

(1) The number of infected messages i has already received from other neighbors is either $\tau - 1$ or $\tau - 2$. (That both of these cases can occur with positive

probability uses the condition that $\sum_{\ell,j} g_{ij}^{\ell} > \tau$, so that there are at least $\tau - 1$ layer-connections from i to others besides j, k.)

(2) At least one of the neighbors j and k is infected.

The conditional probability (given that one is in one of these four cases) that i gets infected can be found in the table below. The top entry in each cell represents the multiplexing scenario and the bottom represents the unmultiplexed case.

	$\tau - 1$	$\tau - 2$
both j, k infected	$q^A + q^B - f(2, \{A, B\})$ $\wedge\vert$ $q^A + q^B - q^A q^B$	$f(2, \{A, B\})$ $\vee\vert$ $q^A q^B$
one of j, k infected	$(q^A + q^B - f(2, \{A, B\}))/2$ $\wedge$ $(q^A + q^B)/2$	$f(2, \{A, B\})/2$ $\vee$ 0

The inequality indicates the direction of which probability is larger. The $\tau - 1$ column (aggregating over both rows which have positively probability) has strictly higher probability for the unmultiplexed case, while the $\tau - 2$ column has strictly higher probability for the multiplexed case. Let ϕ be the probability on the first column and ψ on the second column, and note that the conditional probability of the first row is ρ^2 and the second row is $2\rho(1 - \rho)$. The differences in overall probabilities of infection of the unmultiplexed minus unmultiplexed is then

$$(\psi - \phi) \left[\rho^2 (f(2, \{A, B\}) - q^A q^B) + 2\rho(1 - \rho) f(2, \{A, B\})/2 \right].$$

Given that $f(2, \{A, B\}) - q^A q^B \geq 0$, then this expression has the sign of $(\psi - \phi)$. The proof is then completed by noting that for high enough ρ, q^A, q^B the first column becomes more likely than the second, and for low enough ρ, q^A, q^B the second column becomes more likely than the first.This is where the condition that $(1 + \varepsilon) q^A q^B \geq f(2, \{A, B\}) \geq q^A, q^B$ is invoked. With independent signal transmission across layers, for low enough ρ, q^A, q^B, it is strictly more probable to have fewer than more signals from the connections other than j, k, and thus $\psi - \phi > 0$. These probabilities are continuous in f and so this holds for some $\varepsilon > 0$. The reverse is true for high enough ρ, q^A, q^B. ∎

Proof of Proposition 4: We begin with the case of sufficiently low q^A, q^B and high δ. In that case ρ will also be low (an absolute bound is simply $(q^A + q^B)/\delta$ as that

is a crude upper bound on the infection rate of any given node that always has all neighbors infected and needs only one signal). Then we can invoke Proposition 3 for each connection configuration (noting that there are a finite number of them, taking the min over the $\underline{\rho}$), and then the remaining argument is analogous to the proof of Proposition 2. The reverse holds for the case of sufficiently high q^A, q^B and low δ. ∎

Proposition 5. *The relation $\prec$ is acyclic.*

Proof of Proposition 5 Recall that we denote the set of layers a link ij belong to by $\mathcal{L}_{ij}$. Define the total multiplexity index of a multigraph g as $S_g = \sum_{i>j} |\mathcal{L}_{ij}|^2$.

We show that if $\widehat{g} \prec g$, then $S_g > S_{\widehat{g}}$. By our definition of $\widehat{g} \prec g$, we know that there exist nodes i, j, k and layers ℓ, ℓ' such that $g^\ell_{ik} = 0 = \widehat{g}^\ell_{ij}$ and $\widehat{g}^\ell_{ik} = 1 = g^\ell_{ij}$, all else being equal. We only focus on the contribution of these edges in total multiplexing index since all other links are identical across the two multigraphs. For the multigraph g, this can be represented as $|\mathcal{L}_{ij}|^2 + |\mathcal{L}_{ik}|^2$, while for the less multiplexed graph $\widehat{g}$, the contribution of these edges can be written as $(|\mathcal{L}_{ij}| - 1)^2 + (|\mathcal{L}_{ik}| + 1)^2$. We can then write the difference in total multiplexing between g and $\widehat{g}$ as

$$\begin{aligned} S_g - S_{\widehat{g}} &= |\mathcal{L}_{ij}|^2 + |\mathcal{L}_{ik}|^2 - (|\mathcal{L}_{ij}| - 1)^2 - (|\mathcal{L}_{ik}| + 1)^2 \\ &= |\mathcal{L}_{ij}|^2 + |\mathcal{L}_{ik}|^2 - |\mathcal{L}_{ij}|^2 - |\mathcal{L}_{ik}|^2 - 2 + 2|\mathcal{L}_{ij}| - 2|\mathcal{L}_{ik}| \\ &= 2(|\mathcal{L}_{ij}| - (|\mathcal{L}_{ik}| + 1)) \end{aligned}$$

By $\widehat{g} \prec g$, we know that $|\mathcal{L}_{ik}| < |\mathcal{L}_{ij}| - 1$ (recall that we assumed i and j were linked in at least two layers), hence $S_g > S_{\widehat{g}}$

Now, assume that there exists a cycle such that we have a sequence of multigraphs g_i with $g_1 \prec g_2 \prec g_3 \prec \cdots \prec g_n \prec g_1$. But our proof implies $S_{g_1} < S_{g_2} < S_{g_3} < \cdots < S_{g_n} < S_{g_1}$, which gives us a contradiction. Hence the relation is acyclic. ∎

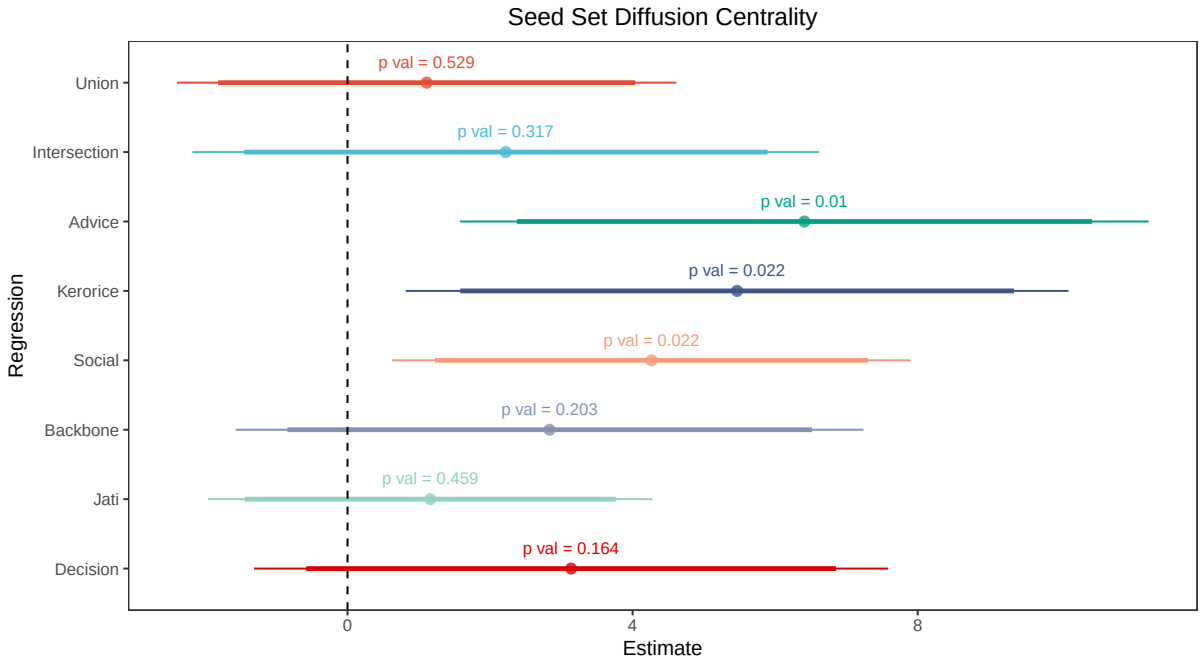
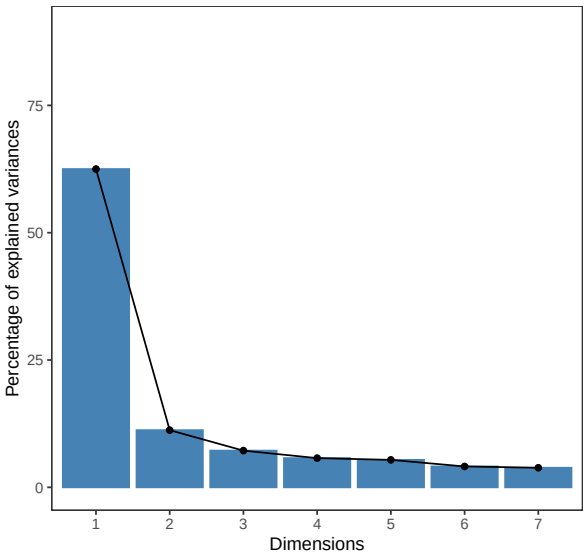
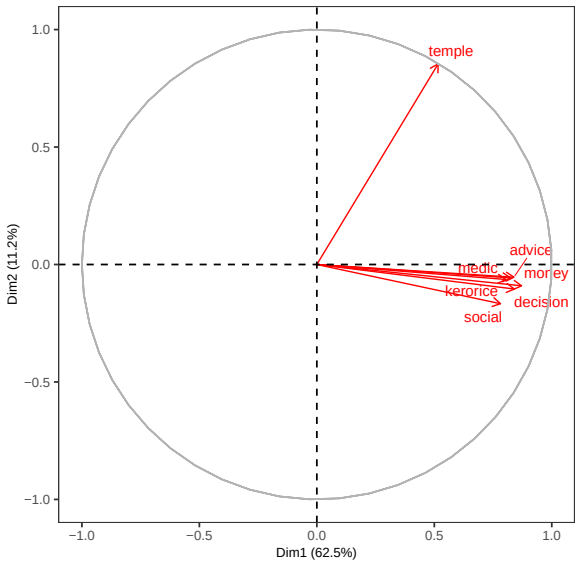


FIGURE S1. OLS Estimates for Impact of Seed Set Diffusion Centrality

A.2. **Supplementary Figures.** Figure S1 plots the results from Table 2. $\hat{\beta}^{\ell}$ and both the 90% and 95% confidence intervals for each of the distinct layers are plotted. Seed centrality in the jati network is not statistically significantly associated with diffusion ($p = 0.459$). Seed centrality in the advice, social, and kerorice networks all are significantly positively associated with diffusion. The point estimates are large, roughly a 59% increase.



(A) Scree Plot: Microfinance Villages



(B) Principal Components: Microfinance Villages

FIGURE S2. Principal Component Analysis with the Temple Layer, but without Geography or Jati layers.

A.3. **Supplementary Tables.**

TABLE S1. Component Loadings: Microfinance Villages

Network	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
social	0.37	-0.04	-0.03	0.18	-0.55	0.69	-0.13	-0.17	-0.07
kerorice	0.38	0.02	0.01	0.07	-0.43	-0.69	-0.38	-0.17	-0.11
money	0.41	0.06	0.02	0.11	0.07	0.01	-0.14	0.72	0.52
advice	0.40	0.08	0.03	0.07	0.51	0.11	-0.22	0.16	-0.70
decision	0.40	0.07	0.02	0.12	0.47	0.03	-0.01	-0.62	0.45
medic	0.39	0.05	0.01	0.07	-0.12	-0.17	0.88	0.04	-0.14
temple	0.24	0.06	0.02	-0.96	-0.05	0.08	-0.02	-0.03	0.03
jati	0.10	-0.66	-0.74	-0.04	0.08	-0.04	0.01	0.02	0.00
distance	-0.07	0.73	-0.68	0.02	-0.05	0.01	-0.02	-0.01	0.00

TABLE S2. Component Loadings: RCT Villages

Network	PC1	PC2	PC3	PC4	PC5
social	0.49	0.03	-0.58	0.52	-0.39
kerorice	0.50	0.04	-0.39	-0.48	0.60
advice	0.50	0.05	0.37	-0.51	-0.59
decision	0.49	0.05	0.61	0.50	0.37
jati	0.09	-1.00	0.02	0.00	0.00

TABLE S3. F-tests for the layers excluding constructed layers

layer	df	R.sq.	F-stat	p-val	F-stat marginal	p-val marginal
advice	1	0.233	20.057	0.000		
jati	2	0.263	2.628	0.110	2.628	0.110
decision	3	0.272	1.728	0.186	0.834	0.365
kerorice	4	0.293	1.768	0.162	1.804	0.184
social	5	0.293	1.306	0.278	0.006	0.938

A.4. **Algorithms.** ..

Algorithm 1: Diffusion Simulation on Multiplexed Networks

Input: Multiplexed network's adjacency matrix $G = \{G^{(1)}, G^{(2)}\}$, transmission probability q, infection threshold τ, recovery probability δ, initial set of infected nodes I_0

Output: Share of infected nodes in steady state

Definitions:

- N: Set of all nodes in the network, $|N| = n$
- S_t: Set of susceptible nodes at time t
- I_t: Set of infected nodes at time t
- $\sigma_{i,t}$: State of node i at time t, where $\sigma_{i,t} \in \{S, I\}$
- $E_{i,t}$: Number of exposures (infections) node i is exposed to at time t

Step 1: Initialize $S_0 = N \setminus I_0$, I_0;

Step 2: **while** $t <$ *1000* **do**

  foreach $i \in S_t$ **do**

    Calculate $E_{i,t} = \sum_{j \in N} \sum_{l=1}^{2} G_{ij}^{(l)} \cdot \mathbb{I}(\sigma_{j,t} = I) \cdot q$;

    if $E_{i,t} \geq \tau$ **then**

      Node i becomes infected: $\sigma_{i,t+1} = I$;

    end

  end

  foreach $i \in I_t$ **do**

    Node i recovers with probability δ: $\sigma_{i,t+1} = S$ with probability δ;

  end

  Update $S_{t+1} = \{i \in N \mid \sigma_{i,t+1} = S\}$;

  Update $I_{t+1} = \{i \in N \mid \sigma_{i,t+1} = I\}$;

  if $abs(\frac{|I_{t+1}|}{n} - \frac{|I_t|}{n}) < 1e-8$ **then**

    break;

  end

end

Step 4: After convergence, run the simulation for an additional 100 iterations to stabilize the results and take the average across these iterations;

Algorithm 2: Multiplexed Network Generation

Input: Three network layers represented as adjacency matrices: A_1, A_2, A_3

Output: Two multiplexed networks M_1 and M_2

Step 1: Use `keroricego`, `visitgo`, and `advise` as the three matrices respectively;

Step 2: Sort A_2 and A_3 based on their average out-degree, in descending order;

Step 3: Prune the network with the higher average out-degree (among A_2 and A_3) to match that of the network with the lower average out-degree. Denote the pruned network as A_2';

Step 4: Generate the first multiplexed network, M_1, by combining the adjacency matrices of A_1 and the unpruned network (either A_2 or A_3, whichever had the lower out-degree);

Step 5: Generate the second multiplexed network, M_2, by combining the adjacency matrices of A_1 and A_2';