

GOSSIP: IDENTIFYING CENTRAL INDIVIDUALS IN A SOCIAL NETWORK

ABHIJIT BANERJEE[†], ARUN G. CHANDRASEKHAR[‡], ESTHER DUFLO[§],
AND MATTHEW O. JACKSON^{*}

ABSTRACT. Can we identify the members of a community who are best-placed to diffuse information simply by asking a random sample of individuals? We show that boundedly-rational individuals can, simply by tracking sources of gossip, identify those who are most central in a network according to “*diffusion centrality*,” which nests other standard centrality measures. Testing this prediction with data from 35 Indian villages, we find that respondents accurately nominate those who are diffusion central (not just those with many friends). Moreover, these nominees are more central in the network than traditional village leaders and geographically central individuals.

JEL CLASSIFICATION CODES: D85, D13, L14, O12, Z13

KEYWORDS: Centrality, Gossip, Networks, Diffusion, Influence, Social Learning

Date: This Version: August 2014.

Ben Golub, Michael Dickstein and participants at various seminars/conferences provided helpful comments. Financial support from the NSF under grants SES-1156182 and SES-1155302, from the AFOSR and DARPA under grant FA9550-12-1-0411, and from ARO MURI under award No. W911NF-12-1-0509 is gratefully acknowledged. We thank Francisco Muñoz for excellent research assistance. We are grateful to CMF at IFMR, Tanay Balantrapu, Gowri Nagraj, and Manaswini Rao.

[†]Department of Economics, MIT.

[‡]Department of Economics, Stanford University.

[§]Department of Economics, MIT.

^{*}Department of Economics, Stanford University; Santa Fe Institute; and CIFAR.

1. INTRODUCTION

“The secret of my influence has always been that it remained secret.” – Salvador Dalí

Knowing who is influential, or central, in a community, can be important for the community’s members as well as for businesses and policymakers. In particular, the extent to which a piece of information diffuses among a population often depends on how central the initially informed are within the network (see Katz and Lazarsfeld (1955); Rogers (1995); Kempe, Kleinberg, and Tardos (2003, 2005); Borgatti (2005); Ballester, Calvó-Armengol, and Zenou (2006); Banerjee, Chandrasekhar, Duflo, and Jackson (2013)). Policymakers and organizations can thus benefit from targeting the right individuals in efforts to effectively spread valuable information.

However, learning who is central in a social network has the potential to be difficult. For policymakers, collecting detailed network data is costly. Even for members of the community, knowing the structure of the network beyond their immediate friends is far from automatic. Do people know how central other people, outside of their immediate circle, are? In this paper, we answer this question theoretically and empirically.

First, we develop a simple model, building upon our previous work (Banerjee, Chandrasekhar, Duflo, and Jackson, 2013), to show that individuals in a network should be able to identify central individuals within their community *without knowing anything about the structure of the network*. Our model is about a process we call gossip, where nodes generate pieces of information that are then stochastically passed from neighbor to neighbor. We assume that individuals who hear the gossip are able to keep count of the number of times each person in the network get mentioned.¹ We show that for any listener in the network, the relative ranking under this count converges over time to the correct ranking of every node’s ability to send information to the rest

¹We use the term “gossip” to refer to the spreading of information about particular people. Our diffusion process is focused on basic information that is not subject to the biases or manipulations that might accompany some “rumors” (e.g., see Bloch, Demange, and Kranton (2014)).

of the network. The specific measure of a node’s ability to send information that we use is given by the “diffusion centrality,” introduced in [Banerjee et al. \(2013\)](#) which answers the question of how widely information from a given node diffuses in a given number of time periods and for a given random per period transmission probability. We relate diffusion centrality to other standard measures of centrality in the literature, proving that it nests three of the most prominent measures: degree centrality at one extreme (if there is just one time period of communication), and eigenvector centrality and Katz-Bonacich centrality at the other extreme (if there are unlimited periods of communication). For intermediate numbers of periods diffusion centrality takes on a wide range of other values.

In other words, by listening and keeping count of how often they hear *about* someone, individuals learn the correct ranking of community members from the point of view of how effectively they can serve as a source of information to the rest of the community.

Second, we use a unique dataset to assess whether this holds empirically. We asked every adult in each of 35 villages to name the person in their village best suited to initiate the spread of information. We combine their answers (which we call their “nominations”) with detailed network data that include maps of a variety of interactions in each of the 35 villages. We show that individuals nominate highly diffusion/eigenvector central people (on average at the 71st percentile of centrality). We also show that the nominations are not simply based on the nominee’s leadership status or geographic position in the village, but are significantly correlated with diffusion centrality even after conditioning on these characteristics, and are better predictors of centrality than these characteristics. This suggests that people understand our questions and are doing more than simply naming traditional leaders or geographically central individuals. Furthermore, there is suggestive evidence that they nominate people who are indeed diffusion central, not only people with many friends.

In sum, our model shows how individuals can learn who are the most central people in their network, and our empirical work suggests that individuals have learned this (although, of course, the data could be consistent with other models of how people choose individuals to nominate). To our knowledge, this

is the first paper to demonstrate that members of communities are able, easily and accurately, to identify highly central people in their community. It is also the first to describe a simple process by which people can learn things about part of their broader network that may intuitively seem outside their ambit.² This is a result of practical importance, since asking people who is the best person to spread information is a much cheaper way to collect this information than collecting data on the entire network.

2. A MODEL OF NETWORK COMMUNICATION

We consider the following model.

2.1. A Network of Individuals. A society of n individuals are connected via a possibly directed³ and weighted network, which has an adjacency matrix $\mathbf{g} \in \{0, 1\}^{n \times n}$. Unless otherwise stated, we take the network \mathbf{g} to be fixed and let $v^{(1)}$ be its first (right-hand) eigenvector, corresponding to the largest eigenvalue λ_1 .⁴ The first eigenvector is nonnegative and real-valued by the Perron-Frobenius Theorem.

Throughout, we assume that the network is (strongly) connected in that there exists a (directed) path from every node to every other node, so that information originating at any node could potentially make its way eventually to any other node.⁵

2.2. Diffusion Centrality. Banerjee et al. (2013) defined a notion of centrality called *diffusion centrality*, based on random information flow through a network according to the following process, which is a variation of a standard diffusion process that underlies many contagion models.⁶

²There are some papers (e.g., Milgram (1967) and Dodds et al. (2003)) that have checked people's abilities to use knowledge of their friends' connections to efficiently route messages to reach distant people, but those rely on local knowledge.

³When defining \mathbf{g} in the directed case, the ij -th entry should indicate that i can tell something to j . In some networks, this may not be reciprocal.

⁴ $v^{(1)}$ is such that $\mathbf{g}v^{(1)} = \lambda_1 v^{(1)}$ where λ_1 is the largest eigenvalue of \mathbf{g} in magnitude.

⁵More generally, everything we say applies to the components of the network.

⁶See Jackson and Yariv (2011) for background and references. A continuous time version of diffusion centrality appears in Lawyer (2014).

A piece of information is initiated at node i and then broadcast outwards from that node. In each period, each informed node informs each of its neighbors of the piece of information and the identity of its original source with a probability $p \in (0, 1]$, independently across neighbors and history. The process operates for T periods, where T is a positive integer.

There are many reasons to allow T to be finite. For instance, a new piece of information may only be “news” for a limited time. Because of boredom, arrival of other news, the topic of conversation may change. By allowing for a variety of T 's, diffusion centrality admits important finite-horizon cases, as well as more extreme cases where agents discuss a topic indefinitely.⁷

Diffusion centrality measures how extensively the information spreads as a function of the initial node. In particular, let

$$\mathbf{H}(\mathbf{g}; p, T) := \sum_{t=1}^T (p\mathbf{g})^t,$$

be the “hearing matrix.” The ij -th entry of \mathbf{H} , $H(\mathbf{g}; p, T)_{ij}$, is the expected number of times, in the first T periods, that j hears about a piece of information originating from i . Diffusion centrality is then defined by

$$DC(\mathbf{g}; p, T) := \mathbf{H}(\mathbf{g}; p, T) \cdot \mathbf{1} = \left(\sum_{t=1}^T (p\mathbf{g})^t \right) \cdot \mathbf{1}.$$

So, $DC(\mathbf{g}; p, T)_i$ is the expected total number of times that some piece of information that originates from i is heard by any of the members of the society during a T -period time interval.⁸ Banerjee et al. (2013) showed that diffusion

⁷Of course this is an approximation and, moreover, different topics may have different T 's. The current model and definition already moves beyond the literature, but even richer models could also be studied.

⁸We note two useful normalizations. One is to compare it to what would happen if $p = 1$ and \mathbf{g} were the complete network \mathbf{g}^c , which produces the maximum possible entry for each ij subject to any T . Thus, each entry of $DC(\mathbf{g}; p, T)$ could be divided through by the corresponding entry of $DC(\mathbf{g}^c; 1, T)$. This produces a measure for which every entry lies between 0 and 1, where 1 corresponds to the maximum possible numbers of expected paths possible in T periods with full probability weight and full connectedness. Another normalization is to compare a given node to the total level for all nodes; that is, to divide all entries of $DC(\mathbf{g}; p, T)$ by $\sum_i DC_i(\mathbf{g}; p, T)$. This normalization tracks how relatively diffusive one node is compared to the average diffusiveness in its society.

centrality was a statistically significant predictor of the spread of information – in that case, about a microfinance program.

Note that this measure allows people to hear the information multiple times from the same person and count those times as distinct reports, so that it is possible for an entry of DC to be more than n . There are several advantages to defining it in this manner. First, although it is possible via simulations to calculate a measure that tracks the expected number of informed nodes and avoids double-counting, this expression is *much* easier to calculate and for many parameter values the two measures are roughly proportional to each other. Second, this version of the measure relates nicely to other standard measures of centrality in the literature, while a measure that adjusts for multiple hearing does not. Third, in a world in which multiple hearings lead to a greater probability of information retention, this count might actually be a better measure of what people learn.⁹

2.3. Diffusion Centrality’s Relation to Other Centrality Measures. It is useful to situate diffusion centrality relative to other prominent measures of centrality in the literature.

Let $d(\mathbf{g})$ denote degree centrality (so $d_i(\mathbf{g}) = \sum_j g_{ij}$). Eigenvector centrality corresponds to $v^{(1)}(\mathbf{g})$: the first eigenvector of \mathbf{g} . Also, let $KB(\mathbf{g}, p)$ denote Katz-Bonacich centrality – defined for $p < 1/\lambda_1$ by:¹⁰

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

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

⁹One could also further enrich the measure by allowing for the forgetting of information, but with three parameters the measure would start to become unwieldy.

¹⁰See (2.7) in Jackson (2008) for additional discussion and background. This was a measure first discussed by Katz, and corresponds to Bonacich’s definition when both of Bonacich’s parameters are set to p .

The difference between the extremes of Katz-Bonacich centrality and eigenvector centrality depends on whether p is sufficiently small so that limited diffusion takes place even in the limit of large T , or whether p is sufficiently large so that the knowledge saturates the network and then it is only relative amounts of saturation that are being measured.¹¹

THEOREM 1.

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

$$DC(\mathbf{g}; p, 1) = pd(\mathbf{g}).$$

- (2) *If $p \geq 1/\lambda_1$, then as $T \rightarrow \infty$ diffusion centrality approximates eigenvector centrality:*

$$\lim_{T \rightarrow \infty} \frac{1}{\sum_{t=1}^T (p\lambda_1)^t} DC(\mathbf{g}; p, T) = \mathbf{v}^{(1)}.$$

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

$$DC(\mathbf{g}; p, \infty) = KB(\mathbf{g}, p); \quad p < 1/\lambda_1.$$

All proofs appear in the Appendix.

The result shows that as T is varied, diffusion centrality nests three of the most prominent and used centrality measures: degree centrality, eigenvector centrality, and Katz-Bonacich centrality. It thus provides a foundation for these measures and spans between them.^{12, 13} Between these extremes, diffusion centrality measures how diffusion process operates for some limited

¹¹Saturation occurs when the entries of $\left(\sum_{t=1}^{\infty} (p\mathbf{g})^t\right) \cdot \mathbf{1}$ diverge (note that in a [strongly] connected network, if one entry diverges, then all entries diverge). Nonetheless, the limit vector is still proportional to a well defined limit vector: the first eigenvector.

¹²This formalizes a result we mention in (Banerjee, Chandrasekhar, Duflo, and Jackson, 2013). An independent formalization appears in Benzi and Klymko (2014).

¹³We also remark on the comparison to another measure: the influence vector that appears in the DeGroot learning model (see, e.g., Golub and Jackson (2010)). That metric captures how influential a node is in a process of social learning. It corresponds to the (left-hand) unit eigenvector of a stochasticized matrix of interactions rather than a raw adjacency matrix. While it might be tempting to use that metric here as well, we note that it is the right conceptual object to use in a process of *repeated averaging* through which individuals update opinions based on averages of their neighbors' opinions. It is thus conceptually different from the diffusion process that we study. One can define a variant of diffusion centrality that works for finite iterations of DeGroot updating.

number of periods. Importantly, as in [Banerjee et al. \(2013\)](#), the behavior in the intermediate ranges can be more relevant for certain diffusion phenomena.

3. RELATING DIFFUSION CENTRALITY TO NETWORK GOSSIP

We now investigate whether and how individuals living in \mathbf{g} end up with knowledge of others' positions in the network that correlates with diffusion centrality without knowing anything about the network structure.

3.1. A Gossip Process. Diffusion centrality considers diffusion from the *sender's* perspective. Let us now consider the same stochastic information diffusion process but from a *receiver's* perspective. Over time, each individual hears information that originates from different sources in the network, and in turn randomly pass that information on. The society discusses each of these pieces of information for T periods. The key point is that there are many such topics of conversation, originating from all of the different individuals in the society, with each topic being passed along for T periods.

For instance, Arun may tell Matt that he has a new car. Matt then may tell Abhijit that “Arun has a new car,” and then Abhijit may tell Esther that “Arun has a new car.” Arun may also have told Ben that he thinks house prices will go up and Ben could have told Esther that “Arun thinks that house prices will go up”. In this model Esther keeps track of the cumulative number of times that bits of information that originated from Arun reach her and compares it with the same number for information that originated from other people. What is crucial therefore is that the news involves the name of the node of origin – in this case “Arun” – and not what the information is about. The first piece of news originating from Arun could be about something he has done (“bought a car”) but the second could just be an opinion (“Arun thinks house prices will go up”). Esther keeps track of the fact that she has heard two different pieces of information originating from Arun.

Recall that

$$\mathbf{H}(\mathbf{g}; p, T) = \sum_{t=1}^T (p\mathbf{g})^t,$$

is such that the ij -th entry, $H(\mathbf{g}; p, T)_{ij}$, is the expected number of times j hears a piece of information originating from i .

We define the *network gossip heard* by node j to be the j -th column of \mathbf{H} ,

$$NG(\mathbf{g}; p, T)_j := H(\mathbf{g}; p, T)_{\cdot j}.$$

Thus, NG_j lists the expected number of times a node j will hear a given piece of news as a function of the node of origin of the information. So, if $NG(\mathbf{g}; p, T)_{ij}$ is twice as high as $NG(\mathbf{g}; p, T)_{kj}$ then j is expected to hear news twice as often that originated at node i compared to node k , presuming equal rates of news originating at i and k .

Note the different perspectives of DC and NG : diffusion centrality tracks how well information spreads from a given node, while network gossip tracks relatively how often a given node hears information from (or about) each of the other nodes.

To end this sub-section two remarks are in order. First, one could allow passing probabilities to differ by information type and pairs of nodes.¹⁴ Indeed, in Banerjee et al. (2013) we allowed different nodes to pass information with different probabilities, and in Banerjee et al. (2014) we allow the probability of communication to depend on the listener’s network position. Although one can enrich the model in many ways to capture specifics of information passing, this simple version captures basic dynamics and relates naturally to centrality measures. Second, we could allow nodes to differ in how frequently they generate new information which is then transmitted to its neighbors. Provided this transmission rate is positively related to nodes’ centralities, the results that we present below still hold (and, in fact, the speed of convergence would be increased).

3.2. Identifying Central Individuals. With this measure of gossip in hand, we show how individuals in a society can estimate who is central simply by

¹⁴We can generalize our setup replacing p with a matrix \mathbf{P} . Now define

$$\mathbf{H}(\mathbf{g}; \mathbf{P}, T) := \left(\sum_{t=1}^T (\mathbf{P} \circ \mathbf{g})^t \right).$$

Here \mathbf{P} can have entries P_{ij} which allow the transmission probabilities to vary by pair. Note that P_{ij} can depend on characteristics of those involved and encode strategic behavior based on the economics being modeled.

counting how often they hear gossip about others. We first show that, on average, individuals' rankings of others according to how much gossip they hear about the others, given by NG_j , are positively correlated with diffusion centrality.

THEOREM 2. *For any $(\mathbf{g}; p, T)$, $\sum_j \text{cov}(DC(\mathbf{g}; p, T), NG(\mathbf{g}; p, T)_j) = \text{var}(DC)$. Thus, in any network with differences in diffusion centrality among individuals, the average covariance between diffusion centrality and network gossip is positive.*

It is important to emphasize that although both measures, NG_i and DC_i , are based on the same sort of information process, they are really two quite different objects. Diffusion centrality is a gauge of a node's ability to send information, while the network gossip measure tracks the reception of information by different nodes. Indeed, the reason that Theorem 2 is only stated for the sum rather than any particular individual j 's network gossip measure is that for small T it is possible that some nodes have not even heard about other nodes, and moreover they might be biased towards their local neighborhoods.¹⁵

Next, we show that if individuals exchange gossip over extended periods of time, every individual in the network is eventually able to *perfectly* rank others' centralities.

THEOREM 3. *If $p \geq 1/\lambda_1$, then as $T \rightarrow \infty$ every individual j 's ranking of others under $NG(\mathbf{g}; p, T)_j$ will be according to the ranking of diffusion centrality, $DC(\mathbf{g}; p, T)$, and hence according to eigenvector centrality, $v^{(1)}$.*

The intuition is that individuals hear (exponentially) more often about those who are more diffusion/eigenvector central, as the number of rounds of communication tends to infinity. As such, in the limit, they assess the rankings

¹⁵ One might conjecture that more central nodes would be better "listeners": for instance, having more accurate rankings than less central listeners after a small number of periods. Although this might happen in some networks, and for many comparisons, it is not guaranteed. None of the centrality measures considered here ensure that a given node, even the most central node, is positioned in a way to "listen" uniformly better than all other less central nodes. Typically, even a most central node might be farther than some less central node from some other important nodes. This can lead a less central node to hear some things before even the most central node, and thus to have a clearer ranking of at least some of the network before the most central node. Thus, for small T , the \sum is important in Theorem 2.

according to diffusion/eigenvector centrality correctly. The result implies that with very little computational ability beyond remembering counts and adding to them, agents come to learn arbitrarily accurately complex measures of others’ centralities, even for people with whom they do not associate.

More sophisticated strategies in which individuals try to infer network topology, could accelerate learning. Nonetheless, this result holds even in a minimal environment wherein individuals do not know the structure of the network and do not tag anything but the topic of conversation (“Arun has a new car”).

The restriction to $p \geq 1/\lambda_1$ is important. For example, as p tends to 0, then individuals hear with vanishing frequency about others in the network, and network distance between people can matter in determining whom they think is the most important.

4. DATA

To investigate the theory presented above, we examine new data, coupled with detailed network data gathered from villages in rural Karnataka (India). The network data consist of network information combined with “gossip” data for 35 villages.

To collect the network data (described in more details in (Banerjee, Chandrasekhar, Duflo, and Jackson, 2013), and publicly available at <http://economics.mit.edu/faculty/eduflo/social>), we surveyed adults regarding with whom they interact.¹⁶ We have data concerning 12 types of interactions for a given survey respondent: (1) whose houses he or she visits, (2) who visit his or her house, (3) his or her relatives in the village, (4) non-relatives who socialize with him or her, (5) who gives him or her medical advice, (6) from whom he or she borrows money, (7) to whom he or she lends money, (8) from whom he or she borrows material goods (e.g., kerosene, rice), (9) to whom he or she lends material goods, (10) from whom he or she gets advice, (11) to whom he or she gives advice, (12) with whom he or she goes to pray (e.g., at a temple, church or mosque). Using these data, we construct one network for each village, at

¹⁶We have network data from 89.14% of the 16,476 households based on interviews with 65% of all adult individuals aged 18-55.

the household level where a link exists between households if any member of any household is linked to any other member of any other household in at least one of the 12 ways.¹⁷ The resulting objects are undirected, unweighted networks at the household level.

To collect the gossip data, we later asked the adults the following two additional questions:

(Loan) *If we want to spread information about a new loan product to everyone in your village, to whom do you suggest we speak?*

(Event) *If we want to spread information to everyone in the village about tickets to a music event, drama, or fair that we would like to organize in your village, to whom should we speak?*

TABLE 1. Summary Statistics

| | mean | sd |
|--|-------|--------|
| households per village | 196 | 61.70 |
| household degree | 17.72 | (9.81) |
| clustering in a household’s neighborhood | 0.29 | (0.16) |
| avg distance between nodes in a village | 2.37 | (0.33) |
| fraction in the giant component | 0.98 | (0.01) |
| is a “leader” | 0.13 | (0.34) |
| nominated someone for loan | 0.48 | (0.16) |
| nominated someone for event | 0.38 | (0.16) |
| was nominated for loan | 0.05 | (0.03) |
| was nominated for event | 0.04 | (0.02) |
| number of nominations received for loan | 0.45 | (3.91) |
| number of nominations received for event | 0.34 | (3.28) |

Notes: for the variables nominated someone for loan (event) and was nominated for loan (event) we present the cross-village standard deviation.

Table 1 provides some summary statistics for our data. The networks are sparse: the average number of households in a village is 196 with a standard

¹⁷Individuals can communicate if they interact in any of the 12 ways, so this is the network of potential communications, and using this network avoids the selection bias associated with data-mining to find the most predictive subnetworks.

deviation of 61.7, while the average degree is 17.7 with a standard deviation of 9.8.

We see that only 5% of households were named in response to the Loan question (and 4% for the Event question) with a cross-village standard deviation of 3%. Thus there is a substantial concordance in who is named as a good initiator of diffusion within a village. In fact, conditional on being nominated, a household was nominated nine times on average.¹⁸ This is perhaps a first indication that the answers may be meaningful.

The fact that less than half of the households were willing to name someone, is itself intriguing. Perhaps people are unwilling to offer an opinion when they are unsure of the answer¹⁹ or possibly were afraid of offending someone, given that they were asked to name just one person.

We label as “leaders” self-help group leaders, shopkeepers, teachers, etc.: 13% of households fall into this category. We use the term as it was defined by the microfinance organization Bharatha Swamukti Samsthe (BSS) as part of their strategy for identifying people to initiate diffusion for their product. BSS approached such social leaders because they were a priori likely to be important in the social learning process and thereby would contribute to more diffusion of microfinance. In our earlier work, [Banerjee et al. \(2013\)](#), we show that there is considerable variation in the centrality of these “leaders” in a network sense, and that this variation predicts the eventual take up of microfinance.

5. EMPIRICAL ANALYSIS

Our theoretical results suggest that people can learn others’ diffusion or eigenvector centralities simply by tracking news they hear through the network. Thus, they should be able to name central individuals when asked whom to use as a “seed” for diffusion. If this is the case, such direct questions can easily be added to standard survey modules to identify central individuals without obtaining detailed network data.

¹⁸We work at the household level, in keeping with [Banerjee et al. \(2013\)](#) who used households as network nodes; a household receives a nomination if any of its members are nominated.

¹⁹See [Alatas et al. \(2014\)](#) for a model that builds on this idea.

5.1. Descriptive Evidence. Figure 1 shows a village network where we highlight both the nominees and the village leaders used by BSS as seeds for micro-finance information, and thus the best guess of this organization of the most highly central people.²⁰ We see in the figure that “leaders” include central households but also some peripheral households. However, nominees appear to be more central and are rarely peripheral. Additionally, nodes that are both leaders and nominated are highly central.

Let us examine the relative network position of nominees in detail. Since the relevant number of periods to compute diffusion centrality is in principle application-dependent, we start with eigenvector centrality, which is a limit as $T \rightarrow \infty$ of diffusion centrality for large enough p (above the inverse of the first eigenvalue). We return to compare what happens with other p and T , below.

Figure 2 builds on the theme presented in Figure 1. Examining the likelihood that a nominated household is highly central, we find that 47% of households that are both nominated and have a leader are within the top 10% of the eigenvector centrality distribution. Furthermore, 23% of the households that are nominated but are not leaders are in the top decile of the centrality distribution.

This contrasts with households that are not nominated, irrespective of whether they have a leader. Only 16% of households that are not nominated but have a leader, and only 7% of households that are not nominated and have no leader, are in the top decile of the eigenvector centrality distribution.²¹ Another way to see this is that the average percentile in the distribution of eigenvector centrality for those nominated for the loan question is 0.71 (and 0.7 for event), while it is 0.62 for the traditional leaders.

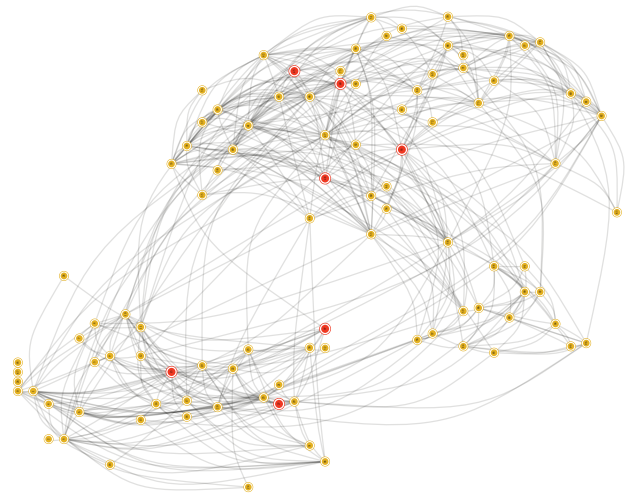
Not only are nominated nodes more central in the network, but more central nodes also receive more nominations, as predicted by the theory. Figure 3 shows that households with higher levels of eigenvector centrality receive more

²⁰This is not a geographic representation of nodes but, rather, a representation using a simple algorithm to visually represent the sub-community structure of the network.

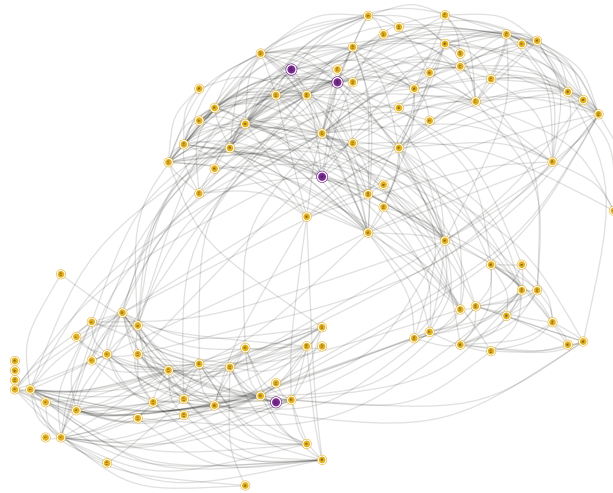
²¹The difference between the 23% of households that are in the top decile given that they are nominated but not leaders and the 16% in the top decile given that they are not nominated but are leaders is significant with a p -value of 0.00 under a Welch test.



(A) Village network with leaders flagged



(B) Village network with nominees flagged



(C) Village network with nominees/leaders flagged

FIGURE 1. A village network with nominees (red), leaders (blue) and those who fulfill both categories (purple) flagged.

nominations on average than households of lower centrality, by a factor of about four when comparing the highest quintile to the lowest.

Figure 4 presents the distribution of nominations as a function of the network distance from a given household. If information did not travel well

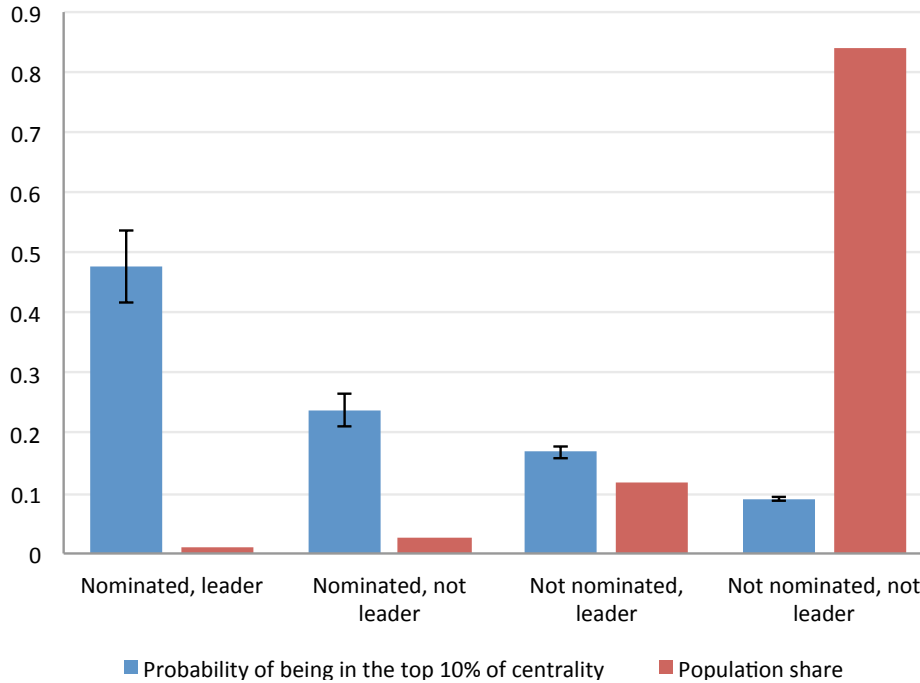


FIGURE 2. The probability that a randomly chosen node with a given classification (whether or not it is nominated under the event question and whether or not it has a village leader) is in the top decile of the eigenvector centrality distribution. 95% confidence intervals are displayed.

through the social network, we might imagine that individuals would only nominate households to whom they are directly connected. Panel A of Figure 4 shows that fewer than 20% of individuals nominate someone within their direct neighborhood. At the same time, over 27% of nominations come from a network distance of at least three or more. Taken together, this suggests that information about centrality does indeed travel through the network.

From Panel A of Figure 4, we also see that, while people do nominate individuals in households who are closer to them than the typical household in the village, respondents generally name people outside of their immediate neighborhoods and sometimes quite far. Moreover, it is important to note that highly central individuals are generally closer to people than the typical household, so, if they did nominate the most central people, this is what we should expect to find.

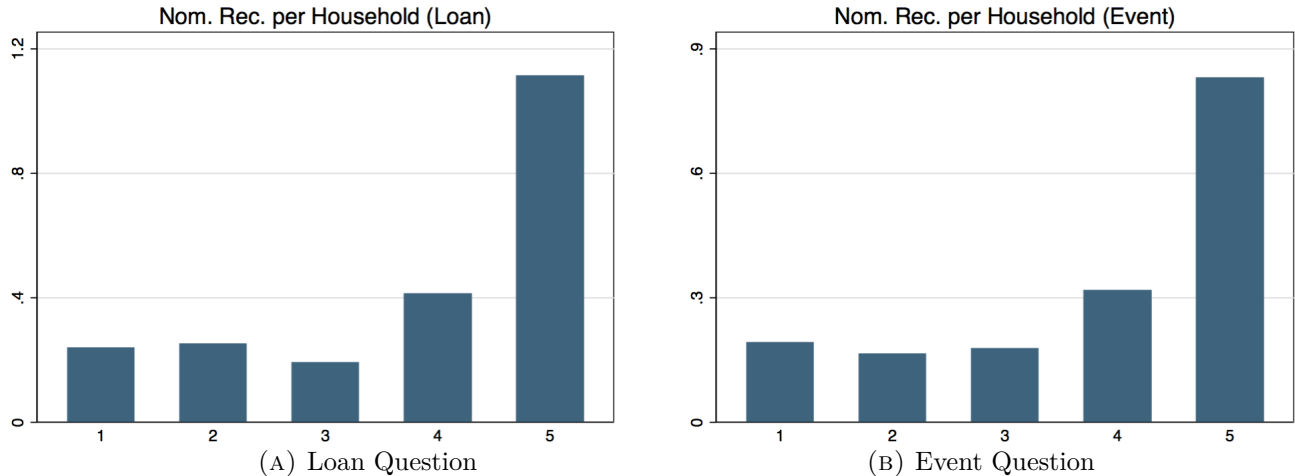


FIGURE 3. The average number of nominations per household vs. the quintile of the households' eigenvector centralities.

It is plausible that individuals may not have good information about who is central in parts of the network that are far from them. However, in Panel B of Figure 4 we see that the average eigenvector centrality percentile of those named at distance 1 is the same as at distance 2 or distance 3 or more. This suggests that individuals have reasonable and comparably accurate information about central individuals in the community who are immediate neighbors or at greater distance from them.

5.2. Regression Analysis. Motivated by this evidence, we present a more systematic analysis of the correlates of nominations, using a discrete choice framework for the decision to nominate someone.

Our theory suggests that if people choose whom to nominate by picking someone whom they hear about most frequently, then diffusion centrality should be a leading predictor of nominations. While the aforementioned results are consistent with this prediction, there are several plausible alternative interpretations. For example, individuals may nominate the person with the most friends, and people with many friends tend to be more diffusion central than those with fewer friends (i.e., diffusion centrality with $T = 1$ and $T > 1$

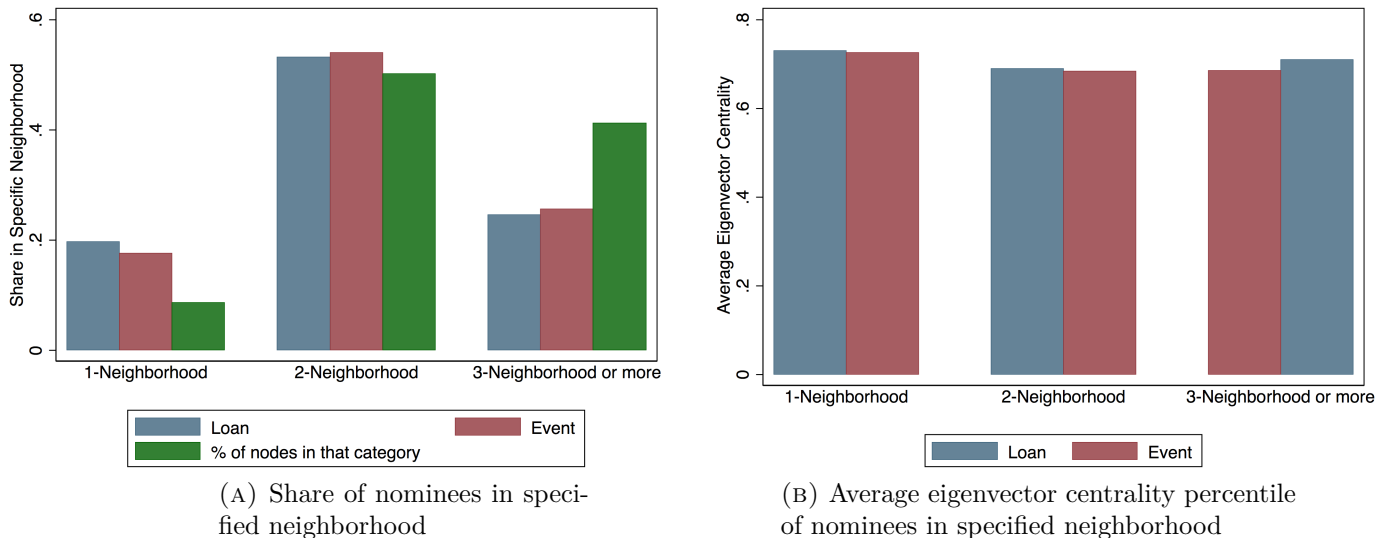


FIGURE 4. Distribution of centralities of nominees

can be positively correlated). Alternatively, it may be that people simply nominate “leaders” within their village, or people who are central geographically, and these also correlate with diffusion/eigenvector centrality. There are indeed a priori reasons to think that leadership status and geography may be good predictors of network centrality, since, as noted in [Banerjee et al. \(2013\)](#), the microfinance organization selected “leaders” precisely because they believed these people would be informationally central. Similarly, previous research has shown that geographic proximity increases the probability of link formation ([Fafchamps and Gubert, 2007](#); [Ambrus et al., 2014](#); [Chandrasekhar and Lewis, 2014](#)) and therefore one might expect geographic data to be a useful predictor of centrality. In addition to leadership data we have detailed GPS coordinates for every household in each village. We include this in our analysis below.²²

²²To operationalize geographic centrality, we use two measures. The first uses the center of mass. We compute the center of mass and then compute the geographic distance for each agent i from the center of mass. Centrality is the inverse of this distance, which we normalize by the standard deviation of this measure by village. The second uses the geographic data to construct an adjacency matrix. We denote the ij entry of this matrix to be $\frac{1}{d(i,j)}$ where $d(\cdot, \cdot)$ is the geographic distance. Given this weighted graph, we compute the eigenvector centrality measure associated with this network. Results are robust to either definition.

Of course, the correlations below are not a “test” that the causal mechanism is indeed gossip, as in our model, but they do rule out that our diffusion centrality measure simply picks out degree centrality, geography or traditional leadership.

To operationalize our analysis using diffusion centrality, we need to first identify p and T . We use the average of the transmission probability parameters estimated in [Banerjee et al. \(2013\)](#), which studies the diffusion of information about a loan, and finds $p = 0.2$. We set $T = 3$, since the average distance between nodes in our network is 2.7 and thus 3 represents the distance to which information travels before it begins to echo backwards. In what follows, we use $DC(0.2, 3)$ when we refer to diffusion centrality.

We estimate a discrete choice model of the decision to nominate an individual. Note that we have large choice sets as there are $n - 1$ possible nominees and n nominators per village network. We model agent i as receiving utility $u_i(j)$ for nominating individual j :

$$u_i(j) = \alpha + \beta'x_j + \gamma'z_j + \mu_v + \epsilon_{ij,v},$$

where x_j is a vector of network centralities for j (eigenvector centrality, $DC(0.2, 3)$, and degree centrality), z_j is a vector of demographic characteristics (e.g., leadership status, geographic position and caste controls), μ_v is a village fixed effect, and $\epsilon_{iv,j}$ is a Type-I extreme value distributed disturbance. For convenience given the large choice sets, we estimate the conditional logit model by an equivalent Poisson regression, where the outcome is the expected number of times an alternative is selected ([Palmgren, 1981](#); [Baker, 1994](#); [Lang, 1996](#); [Guimaraes et al., 2003](#)).

Table 2 presents our estimates of β and γ . First we show that $DC(0.2, 3)$ is a significant driver of an individual nominating another (column 1). A one standard deviation increase in eigenvector centrality is associated with a 0.634 log-point increase in the number of others nominating a household (statistically significant at the 1% level). In column 2 we see that this is robust to controlling for leadership status, geographic centrality of households, as well as village and caste fixed effects: a one standard deviation increase in diffusion centrality is associated with a 0.616 log-point increase in the number of others nominating

TABLE 2. Factors predicting nominations

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|--|---------------------|---------------------|--------------------|-------------------|-------------------|--------------------|--------------------|--------------------|
| $DC(0.2, 3)$ | 0.634*** (0.073) | 0.616*** (0.099) | 0.639** (0.286) | 0.749* (0.399) | 0.982 (0.814) | 0.681* (0.358) | 0.727* (0.427) | 1.208 (1.179) |
| Eigenvector Centrality | | | -0.005 (0.276) | | -0.142 (0.350) | -0.064 (0.349) | | -0.300 (0.594) |
| Degree Centrality | | | | -0.117 (0.405) | -0.208 (0.535) | | -0.113 (0.428) | -0.295 (0.662) |
| Leader | | 0.637** (0.298) | | | | 0.636** (0.298) | 0.640** (0.296) | 0.640** (0.296) |
| Geographic Centrality | | -0.305 (0.216) | | | | -0.303 (0.215) | -0.309 (0.217) | -0.308 (0.217) |
| Observations | 6,466 | 5,733 | 6,466 | 6,466 | 6,466 | 5,733 | 5,733 | 5,733 |
| Village and Caste FE | No | Yes | No | No | No | Yes | Yes | Yes |
| p -value of $\beta_{DC(0.2,3)} = \beta_{\text{Eigenvector}}$ | | | 0.248 | | 0.324 | 0.288 | | 0.389 |
| p -value of $\beta_{DC(0.2,3)} = \beta_{\text{Degree}}$ | | | | 0.279 | 0.373 | | 0.323 | 0.410 |

Notes: The table reports estimate of a Poisson regression where the outcome variable is the expected number of nominations under the loan question. Degree centrality, eigenvector centrality and $DC(0.2, 3)$ are normalized by within-village standard deviation. Standard errors (clustered at the village level) are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

the household (significant at the 1% level). Leadership status does affect the number of nominations as well, as we saw in the descriptive evidence, but the point estimate of the diffusion centrality coefficient hardly changes.

In columns 3-5 we add eigenvector and degree centrality as points of comparison. It seems that it is indeed diffusion centrality that best predicts nominations. A one standard deviation in $DC(0.2, 3)$ corresponds to a 0.639 log-point increase in the number of nominations (statistically significant at the 5% level) as compared to -0.005 log-points when looking at eigenvector centrality (not statistically significant, column 3). Similarly, including both $DC(0.2, 3)$ and degree centrality leaves the coefficient on $DC(0.2, 3)$ mostly unchanged, a 0.749 log-point increase in the number of nominations associated with a one standard deviation increase in centrality, whereas there is no statistically significant association with degree centrality and the point estimate is considerably smaller (-0.117, column 4). However, parameter estimates are noisy and one cannot reject the hypothesis that the coefficients on $DC(0.2, 3)$ and eigenvector centrality as well as $DC(0.2, 3)$ and degree are statistically the same, despite the

difference in the point estimates. The p -values of a test of equality of coefficients are 0.248 and 0.279 in columns 3 and 4, respectively (see the last row of the table). Further, we include all three centrality measures in column 5 and are unable to reject that any individual coefficient is statistically different from zero, although they are jointly significant with a p -value of 0.00). The point estimate of $DC(0.2, 3)$ remains similar to those in column 4, and both degree and eigenvector centrality have negative point estimates. However, due to the colinearity of the three centrality variables, standard errors are more than twice as large than in the previous columns. Columns 6-8 repeat the exercises of columns 3-5, now including leadership status and geographic centrality as extra control variables. Additionally, we add village and household caste fixed effects. As before we find that $DC(0.2, 3)$ enters significantly and with similar effect sizes, unless all three centralities are included, where the estimates become noisy. Moreover, in our pairwise comparisons neither degree nor eigenvector centrality coefficients are large and statistically significant drivers of the number of nominations.

Our results provide suggestive evidence that a key driver of the nomination decision involves diffusion centrality with $T > 1$ although, given the high degree of correlation of these metrics in our sample, we cannot statistically reject that all centrality variables matter equally. The point estimates, however, point towards the diffusion centrality as the most robust factor. Overall, it is clear that individuals are able to accurately name highly central individuals, well beyond using other status variables.

6. CONCLUDING REMARKS

Our model illustrates that it should be easy for even very myopic and non-Bayesian agents, simply by counting, to have an idea as to who is central in their community (according to fairly complex definitions). Motivated by this, we asked villagers to identify central individuals in their village. They do not simply name locally central individuals (the most central among those they know), but actually name ones that are *globally* central within the village. This

suggests that individuals may use simple protocols to learn valuable things about the complex systems within which they are embedded.

Our findings have important policy implications, since such nomination data are easily collected and therefore can be used in a variety of contexts, either on its own, or combined with other easily collected data, to identify who would be a good seed for an information. For instance, if a household is both a nominee and a leader, it has a 47% likelihood of being in the top 10% of the centrality distribution.

It is also worth commenting that our work focuses on the network-based mechanics of communication. In practice, considerations beyond simple network position may determine who the “best” person is to spread information, as other characteristics may affect the quality and impact of communication. An avenue for further research is to investigate whether villagers take such characteristics into account and thus may nominate individuals who are even more successful at diffusing information than the most central individual in the network. Beyond these, the work presented here opens a rich agenda for further research, as one can explore which other aspects of agents’ social environments can be learned in simple ways.

REFERENCES

- ALATAS, V., A. BANERJEE, A. G. CHANDRASEKHAR, R. HANNA, AND B. A. OLKEN (2014): “Network structure and the aggregation of information: Theory and evidence from Indonesia,” *NBER Working Paper*. 19
- AMBRUS, A., M. MOBIUS, AND A. SZEIDL (2014): “Consumption Risk-Sharing in Social Networks,” *American Economic Review*, 104, 149–82. 5.2
- BAKER, S. G. (1994): “The multinomial-Poisson transformation,” *The Statistician*, 495–504. 5.2
- BALLESTER, C., A. CALVÓ-ARMENGOL, AND Y. ZENOU (2006): “Who’s who in networks, wanted: the key player,” *Econometrica*, 74, 1403–1417. 1
- BANERJEE, A., E. BREZA, A. CHANDRASEKHAR, E. DUFLO, AND M. JACKSON (2014): “Come play with me: information diffusion about rival goods,” *mimeo: MIT, Columbia, and Stanford*. 3.1

- BANERJEE, A., A. CHANDRASEKHAR, E. DUFLO, AND M. JACKSON (2013): “Diffusion of Microfinance,” *Science*, 341, DOI: 10.1126/science.1236498, July 26 2013. 1, 2.2, 12, 2.3, 3.1, 4, 4, 18, 5.2
- BENZI, M. AND C. KLYMKO (2014): “A matrix analysis of different centrality measures,” *arXiv:1312.6722v3*. 12
- BLOCH, F., G. DEMANGE, AND R. KRANTON (2014): “Rumors and Social Networks,” *Paris School of Economics, Working paper 2014 - 15*. 1
- BORGATTI, S. P. (2005): “Centrality and network flow,” *Social Networks*, 27, 55 – 71. 1
- CHANDRASEKHAR, A. AND R. LEWIS (2014): “Econometrics of sampled networks,” Stanford working paper. 5.2
- DODDS, P. S., R. MUHAMAD, AND D. J. WATTS (2003): “An Experimental Study of Search in Global Social Networks,” *Science*, 301, 827–829. 2
- FACHAMPS, M. AND F. GUBERT (2007): “The formation of risk sharing networks,” *Journal of Development Economics*, 83, 326–350. 5.2
- GOLUB, B. AND M. JACKSON (2010): “Naive Learning in Social Networks and the Wisdom of Crowds,” *American Economic Journal: Microeconomics*, 2, 112–149. 13
- GUIMARAES, P., O. FIGUEIRDO, AND D. WOODWARD (2003): “A tractable approach to the firm location decision problem,” *Review of Economics and Statistics*, 85, 201–204. 5.2
- JACKSON, M. (2008): *Social and Economic Networks*, Princeton: Princeton University Press. 10
- JACKSON, M. AND L. YARIV (2011): “Diffusion, strategic interaction, and social structure,” *Handbook of Social Economics, San Diego: North Holland, edited by Benhabib, J. and Bisin, A. and Jackson, M.O.* 6
- KATZ, E. AND P. LAZARFELD (1955): *Personal influence: The part played by people in the flow of mass communication*, Free Press, Glencoe, IL. 1
- KEMPE, D., J. KLEINBERG, AND E. TARDOS (2003): “Maximizing the Spread of Influence through a Social Network,” *Proc. 9th Intl. Conf. on Knowledge Discovery and Data Mining*, 137 – 146. 1
- (2005): “Influential Nodes in a Diffusion Model for Social Networks,” *In Proc. 32nd Intl. Colloq. on Automata, Languages and Programming*, 1127

– 1138. 1

- LANG, J. B. (1996): “On the comparison of multinomial and Poisson log-linear models,” *Journal of the Royal Statistical Society. Series B (Methodological)*, 253–266. 5.2
- LAWYER, G. (2014): “Understanding the spreading power of all nodes in a network: a continuous-time perspective,” *arXiv:1405.6707v2*. 6
- MILGRAM, S. (1967): “The small world problem,” *Psychology Today*. 2
- PALMGREN, J. (1981): “The Fisher information matrix for log linear models arguing conditionally on observed explanatory variables,” *Biometrika*, 563–566. 5.2
- ROGERS, E. (1995): *Diffusion of Innovations*, Free Press. 1

APPENDIX A. PROOFS

We prove all of the statements for the case of weighted and directed networks.

Let $v^{(L,k)}$ indicate k -th left-hand side eigenvector of \mathbf{g} and similarly let $v^{(R,k)}$ indicate \mathbf{g} 's k -th right-hand side eigenvector. In the case of undirected networks, $v^{(L,k)} = v^{(R,k)}$. In the case of directed networks, eigenvector $v^{(1)}$ in the main body corresponds to $v^{(R,1)}$.

The following lemma is used in proofs of the theorems.

LEMMA 1. *Consider a positive and diagonalizable \mathbf{g} . Then \mathbf{g} has a unique largest eigenvalue. Moreover, letting $\tilde{\mathbf{g}} = \mathbf{g}/\lambda_1$,*

$$[\tilde{\mathbf{g}}^T]_{ij} \rightarrow_T v_i^{(R,1)} v_j^{(L,1)}.$$

Proof of Lemma 1. The uniqueness of the largest eigenvalue follows from the Perron-Frobenius Theorem (given that g is positive). By diagonalizability we can write \mathbf{g} as

$$\mathbf{g} = \mathbf{V}\Lambda\mathbf{V}^{-1}$$

where Λ is the matrix with eigenvalues on the diagonal (ordered from 1 to n in order of magnitude) and \mathbf{V} is the matrix with columns equal to the right-hand eigenvectors, and \mathbf{V}^{-1} is not only \mathbf{V} 's inverse, but is also the matrix with rows equal to the left-hand eigenvectors. By normalizing $\tilde{\mathbf{g}} = \mathbf{g}/\lambda_1$, it follows that $\tilde{\mathbf{g}} = \mathbf{V}\tilde{\Lambda}\mathbf{V}^{-1}$, where $\tilde{\Lambda}$ is the diagonal matrix such that $\tilde{\Lambda} = \text{diag}\{1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_n\}$, and $\tilde{\lambda}_k = \frac{\lambda_k}{\lambda_1}$. Since the largest eigenvalue of \mathbf{g} is unique, it also follows that $0 \leq |\tilde{\lambda}_k| < 1$ for $k > 1$.

It helps to write the ik -th entry of \mathbf{V} as $v_i^{(R,k)}$, and similarly, the kj -th entry of \mathbf{V}^{-1} is written as $v_j^{(L,k)}$. Therefore, it is straightforward to check that

$$\tilde{\mathbf{g}}_{ij}^T = \sum_{k=1}^n v_i^{(R,k)} v_j^{(L,k)} \tilde{\lambda}_k^T.$$

Separating terms,

$$\tilde{\mathbf{g}}_{ij}^T = v_i^{(R,1)} v_j^{(L,1)} + \sum_{k=2}^n v_i^{(R,k)} v_j^{(L,k)} \tilde{\lambda}_k^T.$$

Thus, given that $\tilde{\lambda}_k^T \rightarrow 0$ for $k \geq 2$, it follows that

$$\tilde{\mathbf{g}}_{ij}^T \rightarrow_{T \rightarrow \infty} v_i^{(R,1)} v_j^{(L,1)},$$

as claimed. ■

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

First, note that in any neighborhood of any nonnegative matrix \mathbf{g} , there exists a positive and diagonalizable matrix \mathbf{g}' . Next, consider any nonnegative \mathbf{g} . If the statement holds for any arbitrarily close positive and diagonalizable \mathbf{g}' , then since $\frac{DC(\mathbf{g}; p, T)}{\sum_{t=1}^T (p\lambda_1)^t}$ is a continuous function (in a neighborhood of a nonnegative and strongly connected \mathbf{g}) as is the first eigenvector, the statement also holds at \mathbf{g} . Thus, it is enough to prove the result for a positive and diagonalizable \mathbf{g} , as in what follows.

In fact, we show the following strengthening of the statement for a positive and diagonalizable \mathbf{g} .

- The ‘tail terms’ approach eigenvector centrality:

$$(\mathbf{g}/\lambda_1)^t \cdot \mathbf{1} \rightarrow v^{(R,1)} \text{ as } t \rightarrow \infty.$$

- If $p > \lambda_1^{-1}$, then

$$\lim_{T \rightarrow \infty} \frac{DC(\mathbf{g}; p, T)}{\sum_{t=1}^T (p\lambda_1)^t} = \lim_{T \rightarrow \infty} \frac{DC(g; p, T)}{\frac{p\lambda_1 - (p\lambda_1)^{T+1}}{1 - (p\lambda_1)}} = v^{(R,1)}.$$

- If $p = \lambda_1^{-1}$, then

$$\lim_{T \rightarrow \infty} \frac{1}{T} DC(\mathbf{g}; \lambda_1^{-1}, T) = v^{(R,1)}.$$

By diagonalizability we write

$$\mathbf{g} = \mathbf{V}\Lambda\mathbf{V}^{-1},$$

and normalizing $\tilde{\mathbf{g}} = \mathbf{g}/\lambda_1$ it follows that $\tilde{\mathbf{g}} = \mathbf{V}\tilde{\Lambda}\mathbf{V}^{-1}$. Normalize the eigenvectors to lie in ℓ_1 , so the entries in each column of \mathbf{V}^{-1} and each row of \mathbf{V}

sum to 1. Then, applying Lemma 1, it follows that $\tilde{\mathbf{g}}^T \cdot \mathbf{1}$ is such that

$$\tilde{\mathbf{g}}^T \cdot \mathbf{1} \rightarrow v^{(R,1)}.$$

This completes the proof of the first statement.

Next, we turn to the sum $\sum_{t=1}^{\infty} \tilde{\mathbf{g}}^t \cdot \mathbf{1}$ and show the second statement for the case where $p = 1/\lambda_1$. It is sufficient to show

$$\lim_{T \rightarrow \infty} \left\| \frac{DC(\mathbf{g}; \lambda_1^{-1}, T)}{T} - v^{(R,1)} \right\| = 0.$$

First, note that as with our calculations above,

$$DC_i(\mathbf{g}; \lambda_1^{-1}, T) = \sum_j \sum_{t=1}^T \sum_k v_i^{(R,k)} v_j^{(L,k)} \tilde{\lambda}_k^t$$

Thus,

$$\begin{aligned} \left| \frac{DC_i(\mathbf{g}; \lambda_1^{-1}, T)}{T} - v_i^{(R,1)} \right| &= \left| \frac{\sum_j \sum_{t=1}^T \sum_{k=1}^n v_i^{(R,k)} v_j^{(L,k)} \tilde{\lambda}_k^t}{T} - v_i^{(R,1)} \right| = \\ &= \left| \frac{1}{T} \sum_j \sum_{t=1}^T \sum_{k=2}^n v_i^{(R,k)} v_j^{(L,k)} \tilde{\lambda}_k^t \right| \leq \frac{1}{T} \sum_{t=1}^T \sum_{k=2}^n 1 \cdot \underbrace{\left| \sum_{j=1}^n v_j^{(L,k)} \right|}_{\leq 1} \cdot |\tilde{\lambda}_k^t| \\ &\leq \frac{n}{T} \sum_{t=1}^T |\tilde{\lambda}_2^t| = \frac{n}{T} \frac{|\tilde{\lambda}_2|}{1 - |\tilde{\lambda}_2|} \left(1 - |\tilde{\lambda}_2|^T \right) \rightarrow 0. \end{aligned}$$

Since the length of the vector (which is n) is unchanging in T , pointwise convergence implies convergence in norm, proving the result.

The final piece repeats the argument for $p > 1/\lambda_1$, but now uses the definition $\tilde{\mathbf{g}} = (p\mathbf{g})$. Then we have $\tilde{\Lambda} = \text{diag}\{\tilde{\lambda}_1, \dots, \tilde{\lambda}_n\}$ with $p\lambda_k = \tilde{\lambda}_k$. We show

$$\lim_{T \rightarrow \infty} \left\| \frac{DC(\mathbf{g}; p, T)}{\sum_{t=1}^T (p\lambda_1)^t} - v^{(R,1)} \right\| = 0.$$

By similar derivations as above,

$$\begin{aligned}
\left| \frac{DC_i(\mathbf{g}; \lambda_1^{-1}, T)}{\sum_{t=1}^T \tilde{\lambda}_1^t} - v_i^{(R,1)} \right| &= \left| \frac{\sum_j \sum_{t=1}^T \sum_{k=1}^n v_i^{(R,k)} v_j^{(L,k)} \tilde{\lambda}_k^t}{\sum_{t=1}^T \tilde{\lambda}_1^t} - v_i^{(R,1)} \right| = \\
&= \left| \frac{1}{\sum_{t=1}^T \tilde{\lambda}_1^t} \sum_j \sum_{t=1}^T \sum_{k=2}^n v_i^{(R,k)} v_j^{(L,k)} \tilde{\lambda}_k^t \right| \\
&\leq \frac{1}{\sum_{t=1}^T \tilde{\lambda}_1^t} \sum_{t=1}^T \sum_{k=2}^n 1 \cdot \left| \sum_{j=1}^n v_j^{(L,k)} \right| \cdot |\tilde{\lambda}_k^t| \\
&\leq \frac{n}{\sum_{t=1}^T \tilde{\lambda}_1^t} \sum_{t=1}^T |\tilde{\lambda}_2^t|.
\end{aligned}$$

Note that this last expression converges to 0 since $\tilde{\lambda}_1 > 1$, and $\tilde{\lambda}_1 > \tilde{\lambda}_2$.²³ which completes the argument. ■

Proof of Theorem 2. Recall that $\mathbf{H} = \sum_{t=1}^T (p\mathbf{g})^t$ and $DC = \left(\sum_{t=1}^T (p\mathbf{g})^t \right) \cdot \mathbf{1}$ and so

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

Additionally,

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

Thus

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

implying

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

²³Note that it is important that $p \geq 1/\lambda_1$ for this claim, since if $p < 1/\lambda_1$, then $p\lambda_1 < 1$. In that case, observe that

$$\frac{\sum_{t=1}^T |\tilde{\lambda}_2|^t}{\sum_{t=1}^T \tilde{\lambda}_1^t} = \frac{\tilde{\lambda}_2}{\tilde{\lambda}_1} \cdot \frac{1 - \tilde{\lambda}_1}{1 - \tilde{\lambda}_2}$$

by the properties of a geometric sum, which is of constant order. Thus, higher order terms ($\tilde{\lambda}_2$, etc.) persistently matter and are not dominated relative to $\sum_{t=1}^T \tilde{\lambda}_1^t$.

which completes the proof. ■

Proof of Theorem 3. Again, we prove the result for a positive diagonalizable \mathbf{g} , noting that it then holds for any (nonnegative) \mathbf{g} .

Again, let \mathbf{g} be written as

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

Also, let $\tilde{\lambda}_k = p\lambda_k$. It then follows that we can write

$$\mathbf{H} = \sum_{t=1}^T (p\mathbf{g})^t = \sum_{t=1}^T \left(\sum_{k=1}^n v_i^{(R,k)} v_j^{(L,k)} \tilde{\lambda}_k^t \right).$$

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

$$\begin{aligned} \mathbf{H}_{\cdot,j} &= \sum_{t=1}^T \left[v^{(R,1)} v_j^{(L,1)} \tilde{\lambda}_1^t + v^{(R,2)} v_j^{(L,2)} \tilde{\lambda}_2^t + O\left(|\tilde{\lambda}_2|^t\right) \right] \\ &= \sum_{t=1}^T \left[v^{(R,1)} v_j^{(L,1)} \tilde{\lambda}_1^t + O\left(|\tilde{\lambda}_2|^t\right) \right] \\ &= v^{(R,1)} v_j^{(L,1)} \sum_{t=1}^T \tilde{\lambda}_1^t + O\left(\sum_{t=1}^T |\tilde{\lambda}_2|^t\right). \end{aligned}$$

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

$$\frac{\mathbf{H}_{\cdot,j}}{\sum_{t=1}^T \tilde{\lambda}_1^t} = v^{(R,1)} v_j^{(L,1)} + O\left(\frac{\sum_{t=1}^T |\tilde{\lambda}_2|^t}{\sum_{t=1}^T \tilde{\lambda}_1^t}\right).$$

Note that the last expression converges to 0 since $\tilde{\lambda}_1 > 1$, and $\tilde{\lambda}_1 > \tilde{\lambda}_2$. Thus,

$$\frac{\mathbf{H}_{\cdot,j}}{\sum_{t=1}^T \tilde{\lambda}_1^t} \rightarrow v^{(R,1)} v_j^{(L,1)}$$

for each j . This completes the proof since each column of \mathbf{H} is proportional to $v^{(R,1)}$ in the limit, and thus has the correct ranking for large enough T .²⁴ Note that the ranking is up to ties, as the ranking of tied entries may vary arbitrarily along the sequence. That is, if $v_i^{(R,1)} = v_\ell^{(R,1)}$, then the ranking that j has over

²⁴The discussion in Footnote 23 clarifies why $p > 1/\lambda_1$ is required for the argument.

i and ℓ could vary arbitrarily with T , but their rankings will be correct relative to any other entries with higher or lower eigenvector centralities. ■