PNAS RESEARCH ARTICLE | COMPUTER SCIENCES SOCIAL SCIENCES

The structure and function of antagonistic ties in village social networks

Amir Ghasemian[a,1] and Nicholas A. Christakis[a]

Edited by Ronald L. Breiger, The University of Arizona, Tucson, AZ; received January 30, 2024; accepted May 15, 2024 by Editorial Board Member Mark Granovetter

Negative or antagonistic relationships are common in human social networks, but they are less often studied than positive or friendly relationships. The existence of a capacity to have and to track antagonistic ties raises the possibility that they may serve a useful function in human groups. Here, we analyze empirical data gathered from 24,770 and 22,513 individuals in 176 rural villages in Honduras in two survey waves 2.5 y apart in order to evaluate the possible relevance of antagonistic relationships for broader network phenomena. We find that the small-world effect is more significant in a positive world with negative ties compared to an otherwise similar hypothetical positive world without them. Additionally, we observe that nodes with more negative ties tend to be located near network bridges, with lower clustering coefficients, higher betweenness centralities, and shorter average distances to other nodes in the network. Positive connections tend to have a more localized distribution, while negative connections are more globally dispersed within the networks. Analysis of the possible impact of such negative ties on dynamic processes reveals that, remarkably, negative connections can facilitate the dissemination of information (including novel information experimentally introduced into these villages) to the same degree as positive connections, and that they can also play a role in mitigating idea polarization within village networks. Antagonistic ties hold considerable importance in shaping the structure and function of social networks.

antagonistic ties | social networks | small-world effect | information diffusion | idea polarization

The overall structure of a social network is determined not only by its positive ties but also by the negative ties within it. But positive and negative relationships have distinct patterns of distribution within networks, and different structural relevance (1–3), and the topological characteristics of friendly and antagonistic ties appear to be fundamentally different (4). While many studies have confirmed classical structural balance theory (e.g., that your friend's enemy is your enemy) (5–10), the real structural impact of antagonistic ties may be even more complex. For instance, antagonists of antagonists may, in fact, be likely to become antagonists themselves rather than friends (2, 3). An incomplete assessment of balance within triads can lead to incorrect conclusions (10, 11). And the evolution of structural balance over time is also important (12).

From an evolutionary perspective, positive interactions, such as cooperation and mutualism, are clearly critical for success, but the potential benefits (if any) of negative relationships are less clear (13). Yet, there are several potential indirect benefits of the capacity for, and existence of, antagonism. For instance, dominance hierarchy plays a crucial role in reducing aggressive encounters and promoting group stability (14, 15). Competition can drive individuals to pursue personal success (16). In the music industry, negative ties with high-status actors may improve sales for low-status actors (17). Or negative relationships might aid in developing certain otherwise useful skills (useful in positive relationships as well) (18), as well as have relevance to personal resilience (19, 20).

Here, we use empirical data from 176 villages in rural Honduras (21) to explore the antagonistic world of social interactions arising from asking 24,770 and 22,513 subjects in two survey waves separated by 2.5 y the question: "Who are the people with whom you do not get along well?" Our primary focus is on the topological structure of these ties and their potentially beneficial role in vital network processes, such as contagion or polarization occurring within the social network as a whole. We are interested in studying the topological location of antagonistic ties within the whole villages and exploring how they might relate to the network structure of the positive world—for example, how established theories of network science, such as the small-world effect, are affected by the

Significance

Our social experience is influenced not only by our positive but also by our negative connections. Using uncommon data from 176 isolated villages in Honduras, we investigate how social network structure and function might be affected by negative ties amid the positive ties of friendship and kinship. We show that having negative ties is associated with people being more peripheral within their subgroups, but closer to other groups within a population, which can have the effect of bringing the whole population closer together. Furthermore, at the collective level, information diffusion is facilitated, and polarization is reduced, by the presence of negative ties. Negative ties can be constructive.

Author affiliations: [a]Yale Institute for Network Science, Yale University, 06511 New Haven, CT

Author contributions: A.G. and N.A.C. designed research; A.G. and N.A.C. performed research; A.G. analyzed data; N.A.C. supervised data collection; and A.G. and N.A.C. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. R.L.B. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: amir.ghasemian@yale.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2401257121/-/DCSupplemental.

Published June 18, 2024.

existence of these ties, and how negative ties influence other metrics of network position otherwise based solely on positive ties.

Furthermore, we are interested in the potential advantages and disadvantages of negative ties with respect to certain important dynamic phenomena in the network. For instance, negative interactions might slow down the spread of contagious diseases in a network that has both positive (attractive) and negative (repulsive) interactions, as compared to a network that is entirely based on positive interactions. Alternatively, negative ties might enhance the dissemination of germs (or information) by locating the nodes linked to more negative connections in the region of networks that are sensitive areas for transmission. Moreover, with respect to information, antagonistic ties within groups might also be associated with a reduction in echo chambers and polarization.

Results

We rely on data collected in a comprehensive sociocentric network study (*SI Appendix*, Table S36) that involved 24,770 individuals at "Wave 1" (W1) and 22,513 individuals at a second wave 2.5 y later ["Wave 2" (W2)] residing in 176 isolated villages in western Honduras (21, 22) (Fig. 1 *A* and *B*). For the sake of simplicity, we only use the giant components of each village network, which encompass an average of 99% of the nodes in each village (comprising a total of 24,621 individuals) during Wave 1 and an average of 98% of the nodes in each village (comprising a total of 22,103 individuals) at Wave 2 (*SI Appendix*, Table S1).

Negative Ties and Network Communities. According to classical balance theory, dynamic systems in the real world progress toward a state with two communities: positive edges are situated within network communities while negative edges are located between them (6). Consequently, community detection, the unsupervised decomposition of a network into groups based on statistical regularities in network connectivity, could assist in identifying the locations of negative ties. We apply a minimum-description-length-based approach [MDL (DC-SBM)] (23), a model-selection-based community detection algorithm that tries to balance the model's complexity with the data's goodness of fit (24) to analyze the positive ties within the network. Contrary to the predictions of classical balance theory, and aligned with more modern assessments of balance theory (12, 25), the results reveal a similar probability of negative ties occurring within communities as compared to between them (no significant difference; see Fig. 1*C* for negative ties and Fig. 1*D* for positive ties). Corresponding results of another community detection method, modularity maximization, are provided in *SI Appendix*, Fig. S1. The results confirm that there are a comparable number of negative ties between and within the communities, in keeping with previous findings suggesting that in order to dislike someone, one first needs to know them ("familiarity breeds contempt") (3).

Next, our analysis centers on defining the "mobility" of the nodes. We measure the change in geodesic distance of the nodes (computed solely through positive ties) from other nodes in their network community, across waves. Using a regression model, we evaluate how the nodes with relatively more negative ties "move" within a network over time relative to their initial communities composed of positive ties. And we analyze both inbound and outbound negative ties.

Our analysis involves $16,017$ individuals shared between the two waves of data (the individuals missing at Wave 2 in our analysis appear to be missing at random with respect to observable topological features and demographic variables; *SI Appendix*, Fig. S4). Using multilevel regression analysis (26) (*Materials and Methods*), we explore the associations between the negative in-degree and out-degree of nodes in Wave 1 and the change in their proximity to their neighbors from Wave 1 at Wave 2.

The multilevel regression model considered in our analysis can be summarized as follows: $\bar{d}_{ij} \sim X_{ij}\beta + u_{j[i]} + \epsilon_i$ and $u_{j[i]} \sim \mathcal{N}(0, \sigma_j^2)$, where $\bar{d}_{ij}$ is the average geodesic distance of node i in village j from the W1's within-community neighbors (in the positive network, as usual) at Wave 2. In order to ensure comparability, the distances have been adjusted by normalizing them according to the diameter of the network (the largest geodesic distance between any pair of nodes in that network). The matrix X_{ij} denotes all characteristics of individual i at village j, including individual i's age, education, gender, religious affiliation, relationship status (ego features); as well as the average age and educational level of neighbors of individual i, and the distributions of gender, religious affiliation, and relationship status for these individuals who are connected to individual i (alter features). Also, $j[i]$ denotes village j that includes individual i. Since the nodes i were at a minimum geodesic distance from their neighbors (distance 1) at Wave 1, it is clear that the distance will necessarily increase at Wave 2.

We find that, over time, the average geodesic distance (as usual, measured solely with positive ties) to their Wave 1 neighbors increases substantially for nodes with higher negative outbound ties, compared to those with fewer negative outbound ties (Fig. 2*A*; see Fig. 2*B* for an illustration; see also *SI Appendix*, Fig. S2A and Table S2 for more details). These results are obtained whether the negative ties are assessed as within or between communities. The relationship of this particular effect with the "movement" of nodes is somewhat stronger than the impact of positive ties between different communities (*SI Appendix*, Fig. S2 *B*, *Top*). The observed contrast in the effects emphasizes the criticality of negative ties in communication with communities beyond one's own.

Nodes with higher levels of within-community "inbound" negative ties at Wave 1 do not move significantly further away compared to those with fewer negative inbound ties. This could be due to the fact that egos may not even be aware of the presence of these inbound ties. However, this effect is larger (and statistically significant) when assessing the geodesic distance in directed networks (*SI Appendix*, Fig. S3). Furthermore, nodes that receive a high number of negative ties from external communities exhibit (on average) less mobility than those with a lower number of such connections (Fig. 2*A* and *SI Appendix*, Figs. S2 and S3).

Network Topology. Given two extreme mechanisms to create two counterfactual worlds—(1) local familiarity (each node is connected to k nearest geographic, i.e., spatial, neighbors), or (2) random familiarity (each node is connected to k random nodes)—we would end up with two extreme worlds of a (k-regular) ring lattice versus an Erdős-Rényi (k-regular) random graph. In a local world with only positive relationships (attractive forces), the nature and extent of the apparent randomness is harder to explain, but it could, in principle, emerge as a result of repulsive forces. In examining real-world social networks, researchers have identified a variety of characteristics, including two key phenomena that distinguish them from random networks: a higher clustering coefficient and a shorter average path length (27). The former characteristic pertains to positive social connections (4), while the latter characteristic bears an

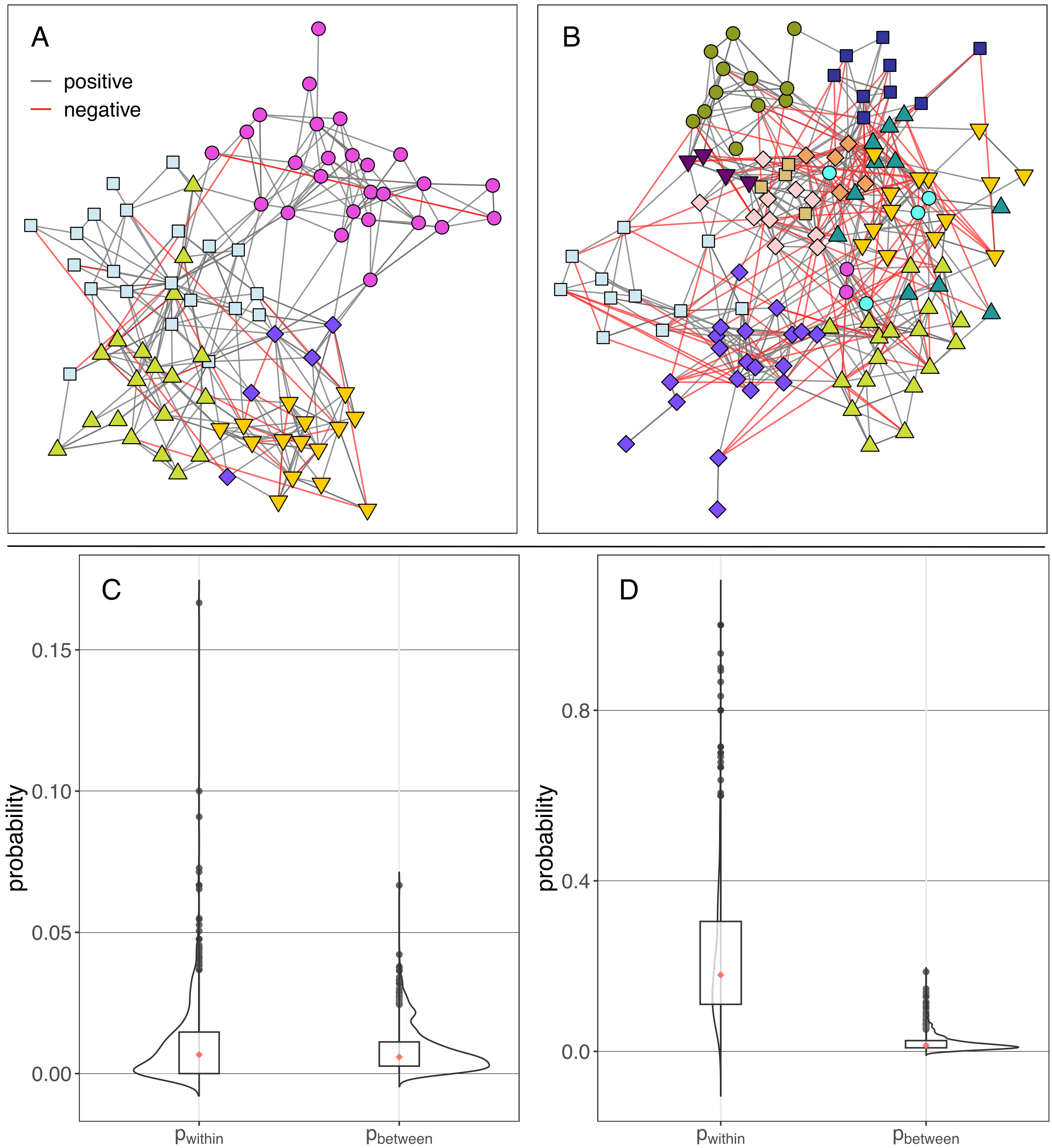

Fig. 1. (*A* and *B*) Two of the village networks in Honduras at Wave 1 (giant component). Positive ties are highlighted by gray edges and negative ties by red edges. Communities are estimated using MDL (DC-SBM) as indicated by the shape and color of the nodes. The likelihood of interactions within a community g can be computed as $p_{\text{within}} = E_{\text{within}}/(n_g(n_g - 1)/2)$, where E_{within} represents the number of edges within the community, and n_g represents the number of nodes in that community. The probability of interactions between two distinct communities g_i and g_j, can be computed as $p_{\text{between}} = E_{\text{between}}/(n_{g_i} n_{g_j})$, where E_{between} is the number of edges between two communities. (*C* and *D*) The plots depict the boxplots of the empirical distribution of negative (*C*) and positive (*D*) interactions within and between communities. The orange diamond represents the median of each probability distribution. As expected, the probability of positive ties within communities is significantly larger than the probability of positive ties between communities. However, the probability of negative ties between and within communities is comparable (no significant difference).

ambiguous nature and could reflect a repulsive force. Hence, next, we investigate the extent to which negative ties might account for the observed shorter average path length.

In order to investigate how antagonistic ties change along with positive ties in a network, we study several underlying (node-specific) topological features, formed solely by positive ties, including: 1) the average geodesic distance from all other nodes in the giant component of the network; 2) the average geodesic distance of a node's neighbors from all other nodes in the giant component of the network; 3) the local clustering coefficient; and 4) the betweenness centrality. [*SI Appendix*, section 3, for three more topological features: 5) the average geodesic distance of a node's neighbors from all other nodes in the giant component of the network when the node itself is removed; 6) the increase in the average geodesic distance of the neighbors from all other nodes in the giant component of the network after the node's

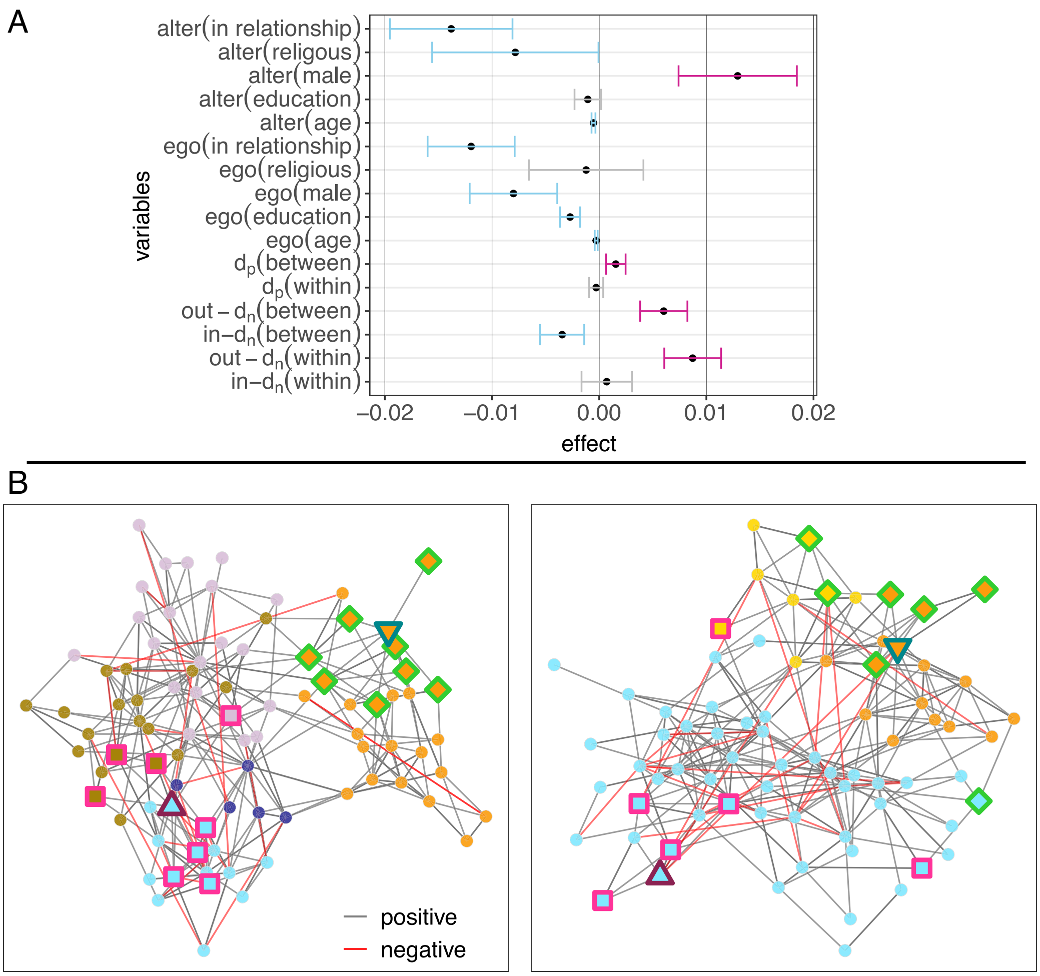

Fig. 2. (*A*) The relationship of negative ties (inbound and outbound) with the "mobility" of nodes (i.e., the change in the geodesic distance of a node to other nodes in its network community across waves, with respect to the W1's within-community neighbors). As expected, a greater number of negative outbound ties within a community will result in a greater geodesic distance from an ego's W1 neighbors at Wave 2. Conversely, the presence of negative inbound ties from outside the community can have the opposite effect. The color blue denotes a negative effect, red signifies a positive effect, and gray represents a statistically insignificant effect. The communities are inferred using MDL (DC-SBM). (*B*) Visualization of the effect of negative ties on mobility. We chose two nodes—represented by two equilateral triangles on both panels—as an illustration. One of these nodes has the negative degree 3 at Wave 1 and its perimeter is colored dark violet, while the other has the negative degree 0 at Wave 1 and its perimeter is colored dark green. Squares represent the W1 neighbors of the node with the larger negative degree with a light violet perimeter color, while diamonds represent the W1 neighbors of the node with the 0 negative degree with a light green perimeter color. In Wave 2, we only include the neighbors that are present in our data. It is clear that the neighbors of the node with the higher negative degree (violet squares) are more dispersed in Wave 2 than the neighbors of the node with the lower negative degree (green diamonds). This diagram is a simple representation of one of the various patterns that can be observed in the data. The interior colors of the nodes represent community membership computed at each wave.

removal; and 7) the average distance from all other nodes across all simple paths of length less than 5 (a simple path is a path with no repeating vertices)].

To compute the association of antagonistic ties with these (positive-world) topological features, we regress these features, measured at the ego level, on the number of antagonistic ties while controlling for covariates such as the size of the network, gender, age, educational level, relationship status, and positive degree. At each village, we normalize all (positive-world) topological features except for the clustering coefficient (*SI Appendix*, section 3). We consider the individual-level fixed effects and village-level random effects using multilevel linear modeling (Fig. 3*A* and *SI Appendix*, Figs. S5 and S6 and Tables S4 and S5). The results indicate that the location of nodes having more negative ties is at the periphery of the communities (at "bridge" locations), and also in regions where we observe smaller clustering coefficients, larger betweenness centralities, and smaller average distances relative to other nodes in the networks (Fig. 3*A*).

In order to gain a deeper understanding of this phenomenon, and the topological location of the nodes with more negative ties, we identified the most central direct (positive) neighbor of each node (i.e., one hop away), measured by the eigenvector centrality of the nodes (in the network of just positive ties). Then, we calculated the likelihood of accessing this node through intermediary nodes (common positive neighbors). The probability of reaching the most important positive neighbor, denoted as k, via common neighbors j between i and k can be expressed as $p_{i\rightarrow j\rightarrow k} = {CN_{ik}}/{d_i^{(+)}}{-1}$, where CN_{ik} is the number of common positive neighbors between i and k and $d_i^{(+)}$ is the positive degree of the original node i. Nodes with higher negative degree tend to have their most significant positive neighbor located outside of their respective community more

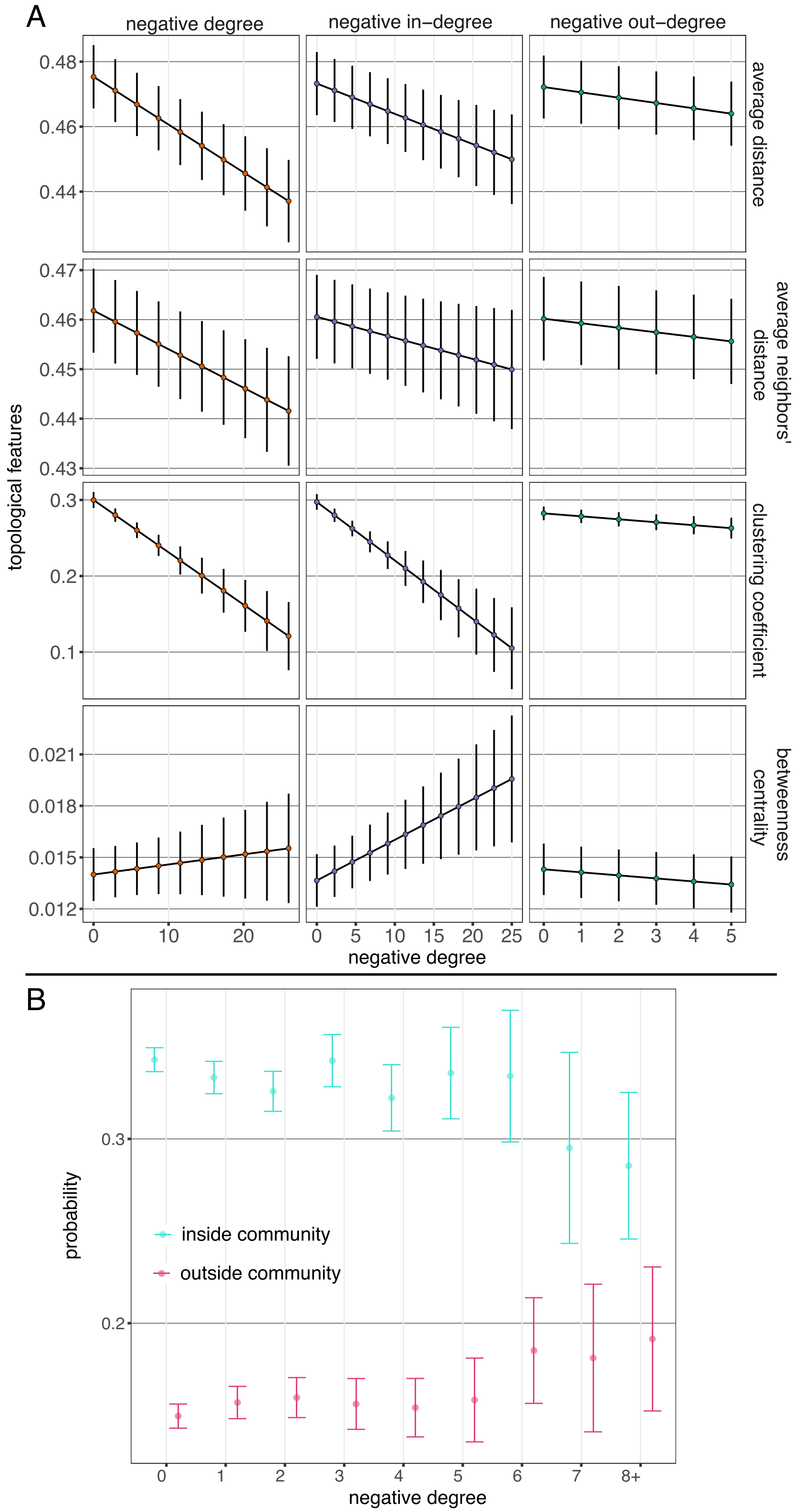

Fig. 3. The relationship of the number of negative ties (undirected, inbound, and outbound) with four topological characteristics: 1) average geodesic distance from all other nodes in the giant component of the network; 2) average geodesic distance of a node's neighbors from all other nodes in the giant component of the network; 3) local clustering coefficient; and 4) betweenness centrality. (*A*) The predictions estimating the average marginal effect for the number of negative ties. As indicated, nodes that have more negative ties are situated on the outskirts of the communities, acting as bridges. These nodes show smaller clustering coefficients, larger betweenness centralities, and smaller average geodesic distances to other nodes in the network. Furthermore, nodes with higher negative connections have neighboring nodes that exhibit a smaller geodesic distance from other nodes in the network, on average. This indicates that the neighbors of nodes with higher negative degrees are more globally dispersed compared to nodes with lower negative degrees that are more locally dispersed. (*B*) The probability of accessing the most central neighbor (with the highest eigenvector centrality) for each node according to the number of negative ties associated with that node (limited to the paths of length 2, i.e., through the common neighbors). The probability that a node will access the most important neighbor outside of its community increases as its negative ties increase, whereas the probability of a node accessing the most important neighbor inside the community decreases as its negative ties increase. As the number of data points decreases from *Left* to *Right*, the width of the CIs increases, reflecting the greater SD. In order to determine the communities in these analyses, MDL (DC-SBM) is used.

frequently when compared to nodes with lower negative degree (see Fig. 3*B* for the MDL communities and *SI Appendix*, Fig. S7 for the modularity communities). In contrast, nodes with a lower negative degree typically have their most significant positive neighbor located within their community.

We also examine the relationship of negative ties with node neighborhood distances, defined solely based on the positive world, in order to understand better the association between negative ties with positive-world topological changes (see Fig. 3*A*, second row). On average, the distances between a node's neighbors and all other nodes decrease slightly as the negative degree of the node increases. This indicates that the neighbors of the nodes with more negative ties are more globally dispersed compared to nodes with fewer negative ties, which are more locally dispersed.

To explore one particular kind of heterogeneity, we also include the associations between triangle motifs (*SI Appendix*, Fig. S8) and the node's topological characteristics (formed solely by positive ties). To not lose statistical power, we considered only triangles formed by undirected links, which are a total of six new parameters (*SI Appendix*, section 3.A). The relationship of the number of negative ties and triangle motifs with all the topological characteristics is provided in *SI Appendix*, Figs. S9 and S10 (*SI Appendix*, Tables S6 and S7). Previously, we observed a correlation between the number of negative connections of a node in a network and its topological location, such as its proximity to bridges. From the results of an analysis of triangle motifs, we learn that this correlation can be partly attributed to a number of heterogeneities stemming from higher-order motifs that characterize such nodes and can be explained by social theories such as balance theory. For example, we observe that the nodes with more negative ties still exhibit smaller clustering coefficients, or nodes with more motifs in the form of n-p-nap (a triangle with a negative and a positive adjacent edge to the node and a positive nonadjacent edge) have smaller average geodesic distance to other nodes in the network.

Finally, to test the hypothesis that negative ties might also contribute to the splitting apart of networks, we model the likelihood of positive within-community interactions to understand how it changes with the probability of within/between-community negative ties and with the probability of between-community positive ties (*SI Appendix*, Fig. S11 and Table S8). The probability of positive ties within a community decreases with the probability of negative ties within the same community. However, the probability of positive ties within a community increases with (both) the probability of positive and negative between-community interactions. These findings indicate that there is a negative correlation between the number of negative connections within a community and the number of positive interactions. Possibly, such an effect could eventually lead to the fracturing of the community, resulting in the formation of multiple groups. However, studying this phenomenon in greater detail requires data with higher time resolution.

Relationship of Antagonistic Ties with Potential Propagations in Networks. Next, we generate synthetic data using SI (and SIS; see *Materials and Methods* and *SI Appendix*, section 4) models (28) to investigate the potential relationship of negative ties with network dynamics. We create data using various parameters that encompass a range of phenomena. These phenomena may include the dissemination of information, the spread of behavior, or the transmission of pathogens within a population of interconnected individuals. In order to achieve a high level of realism in our simulations, we use data based on the actual networks present in the villages. The SI (and SIS; *SI Appendix*) algorithms generate the data by treating (infecting) one node 10,000 times and observing how often any other nodes are "infected" over 30 iterations (through the positive world). By averaging the fraction of infections among all other nodes, and using a multilevel model, we can infer the capacity (vulnerability) of each node to propagate (or receive) a spreading phenomenon (such as an idea or a germ) (29).

Results for a variety of different parameters for the SI model are shown in Fig. 4 (see *SI Appendix*, Tables S9–S12 for more details on SI model, and *SI Appendix*, Fig. S12 and Tables S13–S16 for the SIS model). Our findings indicate that the presence of negative ties (whether inbound or outbound, or ignoring direction) of a node enhances the diffusion emanating from the node, which is consistent with our previous observations on the association between negative ties and the positive-world topological characteristics of a network. Specifically, negative ties have a tendency to isolate the nodes with more negative ties within their network communities, pushing them toward the periphery of the community. This results in a shorter average geodesic distance between these nodes and the rest of the population, impacting their ability to diffuse contagious phenomena. In addition, this is amplified by the earlier observation that the higher the negative degree of the nodes, the more likely they are to be connected to the most central nodes outside of the community (Fig. 3*B*).

There is a positive correlation between positive and negative degree in most of the village networks (Pearson's correlation $= 0.18$, $P < 10^{-16}$). This means that individuals who have a large social circle also tend to have more antagonistic ties. One may therefore ask whether the relationships observed are simply

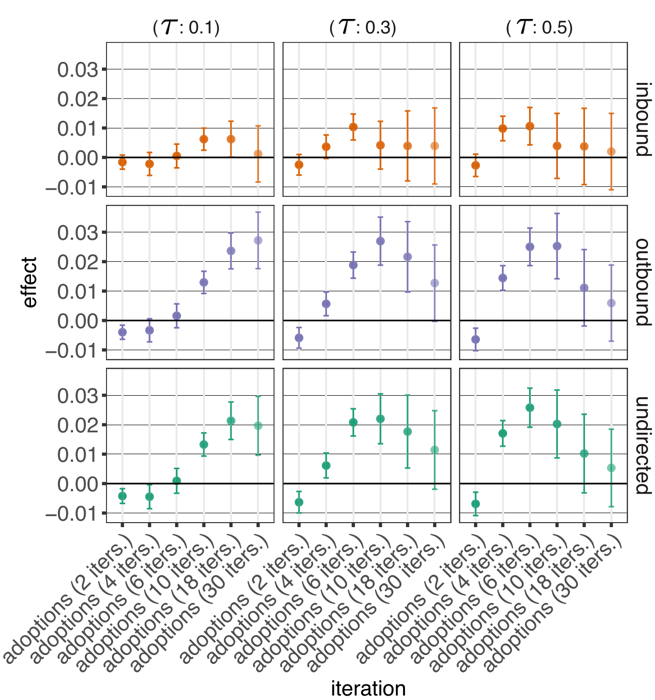

Fig. 4. The CIs of the relationship of the number of negative ties (undirected, inbound, and outbound) with information diffusion and pathogen propagation using the SI model. The presence of negative ties at a node is associated with an augmentation of the process of diffusion emanating from that node, particularly during intermediate iterations. Following multiple iterations in the SI model, the infection eventually reaches a saturation point, which subsequently causes the impact of negative ties to disappear.

a by-product of this correlation. Even though we account for these correlations in all our models, we conducted additional analysis to evaluate the robustness of our findings for the analyses depicted in Fig. 4. By using a null model, we can maintain the correlation while eliminating the impact of the relative location of positive and negative degree. To maintain the positive correlation between positive and negative degree, we rearrange the negative degrees of nodes with the same positive degree. However, this results in a change of the position of the negative degrees in nodes with the same positive degree. That is, in order to keep the correlation fixed, we have to shuffle the negative degrees over the nodes with the same positive degree. The null models we presented do not exhibit the same outcomes, indicating that the observed effects cannot be attributed to positive correlations (*SI Appendix*, Fig. S13). (The corresponding results for the null model of the SIS model are provided in *SI Appendix*, Fig. S14).

We can confirm the accuracy of these simulations by observing a real-life scenario. Our data come from a study aimed at measuring the dissemination of knowledge, attitudes, and practices from randomly targeted households to nontargeted households as part of a randomized controlled trial (RCT) of various social network targeting algorithms (21, 22). Here, we analyze the potential relationship of negative interactions with this diffusion for a particular aspect of the intervention involving the provision of wholly novel information to certain (randomly selected) residents of the villages.

Community health workers had discussions with families about health, and this involved teaching randomly chosen subsets of villagers previously unheard riddles. We analyzed the answers to two riddles to observe how negative relationships are associated with the spread of this exogenously introduced and novel information. We use multilevel logistic regression modeling to determine the association between negative connections and the answers, while also taking into account positive degree, age, gender, education, marital status, and religion (*Materials and Methods*).

Our findings, shown in Fig. 5, indicate that the number of negative ties of a node (whether untargeted or targeted) has the potential to facilitate accurate responses to riddles and enhance the diffusion of information to the node, akin to positive connections (see *SI Appendix*, Table S17 for the effects and CIs for all the covariates). Furthermore, our results suggest that such benefits may extend even to those who were initially uninformed (*SI Appendix*, Fig. S15 and Table S18) (i.e., a person not given the riddle information).

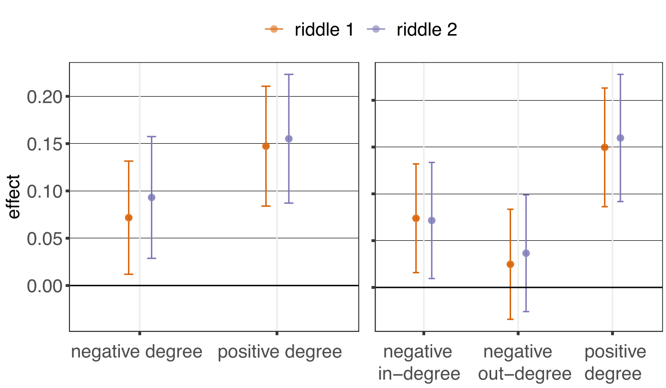

Fig. 5. The relationship of the number of negative ties (undirected edges, inbound, and outbound) with information diffusion. Our analysis involved the evaluation of two logistic regression models. These models aimed to quantify the correlation between the log odds ratio of correct responses to each of two riddles and specific network characteristics and demographic variables. The first model (*Left*) examined the influence of negative and positive degree after controlling for other covariates. The second model (*Right*) extended this examination to include negative in-degree and negative out-degree. Negative ties can have a similar effect as positive ties in terms of increasing the likelihood of providing correct answers to riddles. In other words, overall, negative ties appear to enhance the diffusion of information in a network in a manner that is similar to positive ties. The first riddle (riddle 1) asks, "Dry, dry without a fajero, it falls off quicker, you will see it. What is it?" (umbilical cord), and the second (riddle 2) states, "It seems like it is for the roof, but it's not—it's for diarrhea, you tell me, what is it?" (zinc).

Antagonistic Ties and Idea Polarization. Next, we take advantage of empirical data regarding different beliefs held by individuals. Our analysis focuses on gender-related norms pertaining to how women are treated (such as "In your opinion, in any family, who should decide how the husband's earnings are spent?" "In your opinion, is a husband/companion justified in hitting or beating his wife/companion if she neglects the children?" and "Should the parents of a teenage girl decide if she can join with a partner?").

Considering how negative ties are distributed in a network and how they may influence the network structure, we can formulate hypotheses regarding the potential effect of negative ties within communities on polarization. As mentioned earlier, there were many negative connections within the communities. These connections position the nodes with greater negative degree near the bridges (closer to the rest of the network), possibly contributing to the widespread dissemination of information within the system. We evaluated how these information catalysts (the nodes with more negative ties) potentially contribute to the mitigation of the polarization of beliefs within each village.

We perform an analysis to determine the relationship of the number of negative ties with the diversity of beliefs within each community. After using MDL to infer communities within each village, we compute the diversity of beliefs within each community in each village using Shannon entropy (30). We consider 36 different beliefs gathered from all participants in Wave 1 (*SI Appendix*, Table S19). Using multilevel regression, we can model these diversities and compute the relationship of the number of negative ties inside each community (normalized using the number of nodes inside the community) with the community-level diversities [see *SI Appendix*, Tables S20–S27 (community level)]. We adjust for other variables, including the number of negative ties between the communities, the number of positive ties within and between the communities, the SDs of age and education inside these communities, and the entropies of gender and religion in these communities.

The results for the relationship of the normalized number of negative ties within the communities for these beliefs are provided in *SI Appendix*, Fig. S16 [see *SI Appendix*, section 5 and Tables S20–S27 for more details, *SI Appendix*, Fig. S17 for ego network results (individual level; see also *Materials and Methods*, *Idea Polarization*), and the corresponding individual level *SI Appendix*, Tables S28–S35]. The results reveal that negative connections within a community are associated with a greater diversity of beliefs, and thus a decrease in polarization. The analysis conducted on MDL communities showed significant results in 26 out of 36 beliefs (this seems to apply especially to controversial social norms or behaviors related to gender roles and domestic violence). In other words, the presence of negative connections within a community may promote a wider range of beliefs and diminish structural polarization on more controversial topics.

Choosing Seeds for Diffusion. Finally, it is noteworthy that the capacity for information diffusion between two nodes can vary, even if they have the same number of positive ties, based on the discrepancy in the number of negative ties between them. Using the data generated earlier, i.e., by "infecting" a node 10,000 times, we quantify how often other nodes become infected over a period of 30 iterations. We compare two strategies: choosing a node uniformly among the nodes with the same positive degree or with probability proportionate to its negative degree. The data suggests that selecting a node solely based on its positive degree may not yield optimal results in terms of disseminating information. Instead, it appears that a more effective approach would involve selecting nodes based on both positive and negative degree. On average, a larger negative degree significantly promotes the spread of information among nodes of the same positive degree (see *SI Appendix*, section 6 and Figs. S18 and S19, for selection based on negative degree and negative eigenvector centrality, respectively).

Discussion

We examined positive and negative social interactions in 24,770 and 22,513 people in two consecutive waves in 176 isolated Honduras village social networks to address the potential social role of negative ties. Antagonistic connections are related to network structure and dynamics. The smaller average distance of nodes with a larger negative degree from the rest of a network is reminiscent of random wiring in small-world phenomena (27). That is, negative connections may play a role in fostering pseudorandom wiring with distant components in a network. Relatedly, the nodes with greater negative degree tend to be located near network bridges, where the average distance to the rest of the network is shorter, the betweenness centrality is higher, and the clustering coefficient is lower. Additionally, the neighborhood of nodes with relatively more negative ties exhibits a greater degree of global dispersion in comparison to nodes with lower negative ties.

While there are network measures that compute the centrality of the nodes in information diffusion and other dynamics while taking into account community structure (31, 32), we employed community-agnostic topological measures to demonstrate that our findings are not reliant on any specific community structure. The reason is that different algorithms for community detection can infer different community structures, and there is no definitive ground truth on community structure for real-world networks (24).

These topological factors provide special opportunities for nodes with negative ties in applications like information diffusion. On the one hand, antagonistic connections within a community can push individuals toward the fringes of their group, which can ultimately broaden their connections beyond it and aid the dissemination of information between groups. This can also foster diversity of thought and reduce polarization. On the other hand, an excessive amount of negative ties is likely associated with the presence of multiple communities within a larger population, which may lead to divisions within it. As an example, the growing polarization in the U.S. Congress is not only due to a lack of positive relations between groups but also the negative relations between them (33). Comprehensively studying this would require data with a much higher time resolution and is an area for future research.

Since the presence of negative connections in a network is crucial for determining the positioning of the nodes, their positive connections, and their possible role in propagation dynamics, it may be important to take into account the negative ties for downstream tasks such as influence maximization (34) or proposing targeting strategies (4). Negative ties may also affect community-wide social contagions that interventionists may wish to foster.

It is worth highlighting that investigating face-to-face antagonistic ties is more challenging in comparison to positive ties in part because people may be less aware of who dislikes them than who likes them. Moreover, there may also be measurement errors that can occur due to the use of various name generators to survey negative ties, as well as individuals practicing self-censorship to avoid embarrassment, reduce the risk of social norm violation, and protect their privacy (35, 36). Furthermore, the "effects" discussed here should be apprehended as associations rather than causal relationships. To better understand the causal relationship, we would require data with a more detailed time resolution or a randomized study. This could be a promising direction for future work (for instance, using online experiments).

We considered the spread of social norms through an ordinary kind of contagion process. That is, we focused on the relationship of negative ties with network measures and synthetic dynamics, such as information diffusion in a network, which we also validated using real data (which supported the simplistic assumptions). Nevertheless, it is plausible that social norms not only spread through an ordinary contagion process but also through intentional behaviors, such as peer sanctioning (37), or more complex means; future studies could address the role of intentional behaviors in the process of social norm diffusion.

Just as friendship ties can impose costs (ranging from the demands our friends place on us to the risk of infection that social connections entail) (38), antagonistic ties can offer benefits (ranging from enhancing our access to novel information or reducing our membership in overly siloed groups). There is a distinction, of course, between disliking others or being disliked by them. We have observed that explicit patterns are more discernible in the modeling of disliking others compared to being disliked by them. This is likely primarily because the latter is more likely to remain concealed from the individual. Nevertheless, at the population level, the existence of a capacity for such antagonism has important implications for the overall structure and function of human groups.

Materials and Methods

Our data come from a sociocentric network study of 24,770 people aged 11 to 93 y (with a mean age of 33) at Wave 1 and 22,513 people aged 14 to 95 y (with a mean age of 37) at Wave 2 [labeled as "Wave 3" in the original study (21)] in 176 geographically isolated villages in western Honduras (21). This research was approved by the Yale IRB and by the Honduran Ministry of Health (Protocol # 1506016012), and all participants provided informed consent upon enrolling in the study.

Using this empirical data, we construct 176 binary-directed signed networks (with no multiedges or self-loops). We use three name generators to determine (overlapping) positive ties ("Who do you spend your free time with?" "Who is your closest friend?" and "Who do you discuss personal matters with?") and one specific name generator for negative ties ("Who are the people with whom you do not get along well?"). We consider the weakly connected component of directed networks and the giant component of undirected networks. These components include 99% and 98% of the nodes (on average) at Wave 1 and Wave 2, respectively.

There are 101,997 positive directed interactions—a single positive interaction between ego and alter may represent the result of several positive interactions—compared to 15,776 negative interactions in the 176 villages. These ties include 101,978 positive directed interactions and 15,717 negative interactions in the 176 weakly connected components of these villages (*SI Appendix*, Table S1).

Negative ties have a lower reciprocity than positive relationships. There can be a lack of information flow due to the avoidance of any interaction with the receiver of the antagonistic tie by the sender, which may result in the receiver being unaware of the tie and thus unable to "reciprocate" it (39). In practice, the amount of reciprocity may depend on the question, and the study of undirected networks in Honduras dataset makes sense given that negative ties result from asking "Who are the people with whom you do not get along well?," which constitutes a symmetric relationship.

SI and SIS Model and Propagation in Networks. SI and SIS are two classic compartmental models used in some analyses here to model pathogen and information diffusion (40, 41). The model tracks both the number of "susceptibles" and the number of "infecteds" using simple differential equations. Here, we use the expanded version of the SI and SIS models across a network (28). Through these models, an "infection" is transmitted across an infected-susceptible edge with a probability τ, per iteration. Therefore, any susceptible node i becomes infected with the probability of $1-(1-\tau)^{|\mathcal{I}_{\mathcal{N}_i}|} \approx 1-e^{-\tau|\mathcal{I}_{\mathcal{N}_i}|}$, where $|\mathcal{I}_{\mathcal{N}_i}|$ is the number of infected neighbors of node i. Additionally, in the SIS model, any infected node i is recovered with a constant probability γ at each iteration, resulting in a geometrically distributed infectious period.

The SI and SIS algorithms generate the data by treating (infecting) one node 10,000 times and observing how often any other nodes are "infected" over 30 iterations. By averaging the fraction of infections among all other nodes, we can determine which node has a greater capacity (vulnerability) to propagate (receive) a spreading phenomenon–such as an idea or a germ. Then, in order to determine the impact of negative ties on diffusion, multilevel modeling is used. This involves taking into account the random effects at the village level and making adjustments for any potential confounding factors.

Multilevel Regression. We account for the fact that observations are nested within groups when modeling the relationship between a dependent variable and an independent variable using multilevel regression (26). We use this approach in order to investigate the effect of negative ties on the structure of the social networks, after controlling other factors. We evaluate how the number of negative ties of a node is associated with its topological location in the positive world. Similarly, we study how negative ties are associated with information diffusion and polarization by computing the effect of the number of negative ties with respect to each of these phenomena after controlling other important factors.

We use multilevel regression to model the relationship of negative ties with various topological features, the change in the geodesic location of nodes across different waves, the spread of information (including information introduced exogenously as part of the RCT) or germs, and idea polarization. For each of these cases, we had a quantifiable dependent variable y, which we aimed to model using other covariates. These covariates include age, gender, educational background, relationship status, and topological features like negative tie measures. When evaluating a model at either the individual or community level, certain quantities are evaluated differently depending on the context. For example, when considering the individual level, the negative ties are taken into account, whereas at the community level, the proportion or probability of negative ties is given due consideration. A multilevel regression model at the individual level may be summarized as $y_{ij} \sim X_{ij}\beta + \epsilon_i + u_{j[i]}$ and $u_{j[i]} \sim \mathcal{N}(0, \sigma_j^2)$ to compute the negative ties' effects. The X_{ij} represents all of the characteristics of the individual i at village j, such as its number of negative and positive ties, and the individual's age, gender, educational background, and relationship status (i.e., a binary variable of being in a marital relationship). Also, $j[i]$ denotes village j that includes individual i. For the community level, we have $y_{gj} \sim X_{gj}\beta + u_{j[g]} + \epsilon_g$ and $u_{j[g]} \sim \mathcal{N}(0, \sigma_j^2)$, where X_{gj} represents all of the characteristics of the community g in village j (see next subsection for an example).

In our analysis of the effects of negative ties on information diffusion, we examined the relationship between the number of negative ties and the probability of correctly answering exogenously introduced novel riddles. Our aim is to determine how this relationship can impact information diffusion. To this end, we use a multilevel logistic regression summarized as $\mathrm{logit}(p_{ijk}) \sim X_{ijk}\beta + \epsilon_i + u_{j[i]} + v_{k[j]}$ and $u_{j[i]} \sim \mathcal{N}(0, \sigma_j^2)$ and $v_{k[j]} \sim \mathcal{N}(0, \sigma_k^2)$ to compute the negative ties' effects, where the log odds ratio has been regressed on both the demographic variables and the topological features of the network. Our model accounts for the hierarchical structure of the data, with individuals hierarchically nested within households (since the individuals inside a household are targeted together) and households nested within villages. This approach provides us with the ability to consider the potential random effects that may exist at both the household and village levels, thus reflecting the intrahousehold and intravillage correlations.

Idea Polarization. In order to quantify idea polarization, we measure the diversity of beliefs at the level of the communities (community level) and of the ego networks (individual level). We examine the polarization using the gender norm questions (*SI Appendix*, Table S19) from surveys collected in the villages (21). We quantify the diversity of beliefs within each community or in the neighborhood of each individual using Shannon entropy (30). This measure helps us assess the diversity of beliefs for each of the questions, which is inversely related to polarization.

For computing these effects, we use a multilevel regression model summarized as $\mathrm{polarization}_{gj} \sim X_{gj}\beta + u_{j[g]} + \epsilon_g$ and $u_{j[g]} \sim \mathcal{N}(0, \sigma_j^2)$. The X_{gj} represents the characteristics of community g within village j, including its number of positive and negative ties within that community, its number of positive and negative edges inward and outward of the community, its age and religion entropies, and its variance in age and educational background within the community.

Data, Materials, and Software Availability. The data used in this work are not publicly available given mandated commitments to the research participants and the sensitive nature of the health and social data in these small communities that could potentially allow decryption or individual identification, but data may be available from the senior author on reasonable request and subject to a DUA and establishing a secure server account. Illustrative data from this cohort have also previously been released. The network data for a sample of 22 signed villages are provided at https://github.com/Aghasemian/EnmityParadox (4).

ACKNOWLEDGMENTS. This work was supported in part by the NSF under Grant No. 2030859 to the Computing Research Association for the CIFellows Project (A.G.), and also by the Bill and Melinda Gates Foundation, with additional support from the NOMIS Foundation (Switzerland) and R01AG062668 from the National Institute on Aging.

1. J. Leskovec, D. Huttenlocher, J. Kleinberg, "Predicting positive and negative links in online social networks" in *Proceedings of the 19th International Conference on World Wide Web* (ACM, New York, NY, 2010), pp. 641–650.
2. J. Leskovec, D. Huttenlocher, J. Kleinberg, "Signed networks in social media" in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, New York, NY, 2010), pp. 1361–1370.
3. A. Isakov, J. H. Fowler, E. M. Airoldi, N. A. Christakis, The structure of negative social ties in rural village networks. *Sociol. Sci.* **6**, 197–218 (2019).
4. A. Ghasemian, N. Christakis, The enmity paradox. *Sci. Rep.* **13**, 20040 (2023).
5. F. Heider, *The Psychology of Interpersonal Relations* (Psychology Press, 1982).
6. D. Cartwright, F. Harary, Structural balance: A generalization of Heider's theory. *Psychol. Rev.* **63**, 277 (1956).
7. G. Facchetti, G. Iacono, C. Altafini, Computing global structural balance in large-scale signed social networks. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20953–20958 (2011).
8. E. Estrada, M. Benzi, Walk-based measure of balance in signed networks: Detecting lack of balance in social networks. *Phys. Rev. E* **90**, 042802 (2014).
9. A. Kirkley, G. T. Cantwell, M. E. Newman, Balance in signed networks. *Phys. Rev. E* **99**, 012320 (2019).
10. E. Estrada, Rethinking structural balance in signed social networks. *Discret. Appl. Math.* **268**, 70–90 (2019).
11. D. Feng, R. Altmeyer, D. Stafford, N. A. Christakis, H. H. Zhou, Testing for balance in social networks. *J. Am. Stat. Assoc.* **117**, 156–174 (2022).
12. C. M. Rawlings, N. E. Friedkin, The structural balance theory of sentiment networks: Elaboration and test. *Am. J. Sociol.* **123**, 510–548 (2017).

13. N. A. Christakis, *Blueprint: The Evolutionary Origins of A Good Society* (Hachette UK, 2019).
14. E. A. Tibbetts, J. Pardo-Sanchez, C. Weise, The establishment and maintenance of dominance hierarchies. *Philos. Trans. R. Soc. B* **377**, 20200450 (2022).
15. J. C. Flack, M. Girvan, F. B. De Waal, D. C. Krakauer, Policing stabilizes construction of social niches in primates. *Nature* **439**, 426–429 (2006).
16. B. C. DiMenichi, E. Tricomi, The power of competition: Effects of social motivation on attention, sustained physical effort, and learning. *Front. Psychol.* **6**, 1282 (2015).
17. D. S. Halgin, S. P. Borgatti, Z. Huang, Prismatic effects of negative ties. *Soc. Netw.* **60**, 26–33 (2020).
18. A. R. Overton, A. C. Lowry, Conflict management: Difficult conversations with difficult people. *Clin. Colon Rectal Surg.* **26**, 259–264 (2013).
19. R. G. Tedeschi, L. G. Calhoun, The posttraumatic growth inventory: Measuring the positive legacy of trauma. *J. Trauma. Stress* **9**, 455–471 (1996).
20. S. Maitlis, Posttraumatic growth at work. *Annu. Rev. Organ. Psychol. Organ. Behav.* **7**, 395–419 (2020).
21. H. B. Shakya *et al*., Exploiting social influence to magnify population-level behaviour change in maternal and child health: Study protocol for a randomised controlled trial of network targeting algorithms in rural Honduras. *BMJ open* **7**, e012996 (2017).
22. E. M. Airoldi, N. A. Christakis, Induction of social contagion for diverse outcomes in structured experiments in isolated villages. *Science* **384**, eadi5147 (2024).
23. T. P. Peixoto, Parsimonious module inference in large networks. *Phys. Rev. Lett.* **110**, 148701 (2013).
24. A. Ghasemian, H. Hosseinmardi, A. Clauset, Evaluating overfit and underfit in models of network community structure. *IEEE Trans. Knowl. Data Eng.* **32**, 1722–1735 (2019).
25. P. Doreian, A. Mrvar, Partitioning signed social networks. *Soc. Netw.* **31**, 1–11 (2009).
26. A. Gelman, J. Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models* (Cambridge University Press, 2006).
27. D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
28. O. N. Bjørnstad, *Epidemics: Models and Data Using R* (Springer Nature, 2022).
29. S. V. Shridhar, M. Alexander, N. A. Christakis, Characterizing super-spreaders using population-level weighted social networks in rural communities. *Philos. Trans. R. Soc. A* **380**, 20210123 (2022).
30. C. E. Shannon, A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
31. C. Blöcker, J. C. Nieves, M. Rosvall, Map equation centrality: Community-aware centrality based on the map equation. *Appl. Netw. Sci.* **7**, 56 (2022).
32. S. Rajeh, M. Savonnet, E. Leclercq, H. Cherifi, Characterizing the interactions between classical and community-aware centrality measures in complex networks. *Sci. Rep.* **11**, 10088 (2021).
33. Z. P. Neal, A sign of the times? weak and strong polarization in the us congress, 1973–2016. *Soc. Netw.* **60**, 103–112 (2020).
34. E. Tardos, D. Kempe, J. Kleinberg, "Maximizing the spread of influence in a social network" in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, New York, NY, 2003), pp. 137–146.
35. R. Tourangeau, L. J. Rips, K. Rasinski, *The Psychology of Survey Response* (Cambridge University Press, 2000).
36. P. J. Shoemaker, M. Eichholz, E. A. Skewes, Item nonresponse: Distinguishing between don't know and refuse. *Int. J. Public Opin. Res.* **14**, 193–201 (2002).
37. J. A. Kitts, Social influence and the emergence of norms amid ties of amity and enmity. *Simul. Model. Pract. Theory* **14**, 407–422 (2006).
38. K. P. Smith, N. A. Christakis, Social networks and health. *Annu. Rev. Soc.* **34**, 405–429 (2008).
39. N. Harrigan, J. Yap, Avoidance in negative ties: Inhibiting closure, reciprocity, and homophily. *Soc. Netw.* **48**, 126–141 (2017).
40. R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200 (2001).
41. R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925 (2015).